

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)**SciVerse ScienceDirect**

J. Math. Pures Appl. 98 (2012) 367–389

JOURNAL  
DE  
**MATHÉMATIQUES**

PURES ET APPLIQUÉES

[www.elsevier.com/locate/matpur](http://www.elsevier.com/locate/matpur)

# Unitarity of the KZ/Hitchin connection on conformal blocks in genus 0 for arbitrary Lie algebras <sup>☆</sup>

Prakash Belkale

*Department of Mathematics, UNC-Chapel Hill, CB #3250, Phillips Hall, Chapel Hill, NC 27599, United States*

Received 9 August 2011

Available online 3 March 2012

## Abstract

We prove that the vector bundles of conformal blocks, on suitable moduli spaces of genus zero curves with marked points, for arbitrary simple Lie algebras and arbitrary integral levels, carry unitary metrics of geometric origin which are preserved by the Knizhnik–Zamolodchikov/Hitchin connection (as conjectured by Gawedzki et al., 1991, in [7]). Our proof builds upon the work of Ramadas (2009) [22] who proved this unitarity statement in the case of the Lie algebra  $\mathfrak{sl}_2$  (and genus 0).

© 2012 Elsevier Masson SAS. All rights reserved.

## Résumé

On démontre que les fibrés vectoriels des blocs conformes sur les espaces des modules convenables des courbes de genre zéro avec des points marqués, associés à une algèbre de Lie simple arbitraire et un niveau intégral arbitraire, portent des métriques unitaires d'origine géométrique qui sont préservées par la connexion de Knizhnik–Zamolodchikov/Hitchin (comme cela a été conjecturé par Gawedzki et al., 1991, dans [7]). Notre article s'appuie sur les résultats de Ramadas (2009) [22] qui a démontré cet énoncé dans le cas de l'algèbre de Lie  $\mathfrak{sl}_2$  (et genre zéro).

© 2012 Elsevier Masson SAS. All rights reserved.

MSC: primary 17B67, 14H60; secondary 32G34, 81T40

Keywords: KZ equations; Hitchin connection; Conformal blocks; Non-abelian theta functions

## 1. Introduction

Consider a finite dimensional simple Lie algebra  $\mathfrak{g}$ , a non-negative integer  $k$  called the level and an  $N$ -tuple  $\vec{\lambda} = (\lambda_1, \dots, \lambda_N)$  of dominant weights of  $\mathfrak{g}$  of level  $k$ . The mathematical theory of Tsuchiya and Kanie [25] and Tsuchiya, Ueno, and Yamada [26,18], associates to this data a vector bundle  $\mathcal{V} = \mathcal{V}_{\vec{\lambda},k}$  on  $\overline{\mathcal{M}}_{g,N}$ , the moduli stack of stable  $N$ -pointed curves of genus  $g$ .

On the open part  $\mathcal{M}_{g,N}$  of smooth pointed curves,  $\mathcal{V}$  carries a flat projective connection  $\nabla$ , which is the restriction of a suitable Knizhnik–Zamolodchikov (KZ) connection when  $g = 0$ . The WZW connection [26] generalizes the KZ connection to all genera.

<sup>☆</sup> Partially supported by NSF grant DMS-0901249.

E-mail address: [belkale@email.unc.edu](mailto:belkale@email.unc.edu).

The fibers of  $\mathcal{V}$  on  $\mathfrak{M}_{g,N}$  can also be described in terms of sections of natural line bundles on suitable moduli stacks of parabolic principal bundles on  $N$ -pointed curves of genus  $g$ . These sections generalize classical theta functions, and are hence called non-abelian or generalized theta functions (see the survey [24]). The connection on  $\mathcal{V}$  was described from the above algebro-geometric point of view by Hitchin [12].

A basic conjecture in the subject, with origins in physics, is that  $\mathcal{V}$  carries a projective unitary metric which is preserved (projectively) by the connection  $\nabla$ . It was pointed out to us by A. Kirillov [17] that the combined work of Kirillov and Wenzl [15,16,29] (Theorem 10.9 in [16] and Theorem 3.7 in [29]) implies this conjecture for all genera including genus 0. The construction of the unitary metric via this approach (which uses the theory of quantum groups at roots of unity [14]) is not explicit and is not a priori of Hodge-theoretic origin.

In the 90's Gawedzki and collaborators [11,7,10] proposed a conjectural, explicit construction of the unitary metric via integration of the Schechtman–Varchenko forms [23] (see [23,2,21,28] and the references therein for various applications of these beautiful forms). Recently the case  $\mathfrak{g} = \mathfrak{sl}_2$  and genus 0 of Gawedzki's proposal was rigorously proved by Ramadas [22]. In this paper, following Ramadas' general strategy, we prove the (geometric) unitarity conjecture, as proposed by Gawedzki et al. [7] for arbitrary simple Lie algebras  $\mathfrak{g}$  in genus 0 (this does not follow from the Kirillov–Wenzl approach to unitarity). As in Ramadas' work, the unitary metric is obtained by realizing the bundle of conformal blocks inside a Gauss–Manin system of cohomology of smooth projective varieties. One of the consequences is the production of a large collection of unitary local subsystems of multi-dimensional hypergeometric local systems.

**Theorem 1.1.** *The KZ/Hitchin connection on bundles of conformal blocks over configuration spaces of distinct points on  $\mathbb{A}^1$  is unitary, with the unitary metric of geometric origin, for any simple Lie algebra  $\mathfrak{g}$  and any integral level  $k$ .*

An algebro-geometric strategy for carrying out Gawedzki's proposal was given in [22]. One views conformal blocks as sections of line bundles on moduli of parabolic bundles on  $\mathbb{P}^1$ . In this setting, according to Ramadas, one should first find derivatives of theta functions on Harder–Narasimhan strata, which we take as a suggestion to look at correlation functions. The main steps in Ramadas' proof [22] can be described as follows (here  $g = 0$  and  $\mathfrak{g} = \mathfrak{sl}_2$ ).

- (1) Fix  $N$  marked points  $z_1, \dots, z_N$  on the curve  $C = \mathbb{P}^1$ . The fiber  $V$  of  $\mathcal{V}$  over this marked curve is isomorphic to the space of sections of a natural line bundle on a suitable moduli stack of parabolic principal bundles on  $C$  (with parabolic structure at the given marked points). Any global section of this line bundle necessarily vanishes on any Harder–Narasimhan stratum (which corresponds to non-semistable parabolic bundles). Ramadas' first step is to consider a “Harder–Narasimhan trace”, which is a suitable higher order derivative, at such strata. These considerations lead him to a map from  $V$  to the space of top-degree differential forms on an affine variety  $X$  that depends upon the marked curve (actually  $X$  is an open subset of  $C^M$  for a suitable  $M$ ).
- (2) Ramadas proves (geometrically) a key vanishing statement for such differential forms on partial diagonals. He then lifts these differential forms to a finite cover  $Y$  of  $X$ , multiplies them by a “master function” and shows, using the vanishing statement, that the resulting differential forms extend to any smooth compactification  $\bar{Y}$  of  $Y$ .
- (3) We hence obtain an injective map  $V \rightarrow H^{\dim Y}(\bar{Y}, \mathbb{C})$ . Ramadas then proves that this map is flat for the connections as the marked curve varies (and where we consider the Gauss–Manin connection on the vector bundle formed by the spaces  $H^{\dim Y}(\bar{Y}, \mathbb{C})$ ). To prove the above flatness assertion, it suffices to prove that the composite  $V \rightarrow H^{\dim Y}(Y, \mathbb{C})$  is flat, which is verified by an explicit calculation.
- (4) We are now in a curious situation, with a flat map from  $\mathcal{V}$  to a Gauss–Manin system (of cohomology of smooth projective varieties) whose image is contained in the space of top-degree algebraic forms. It follows that the canonical polarization on the Gauss–Manin system restricts to give the desired metric on  $\mathcal{V}$ .

### 1.1. An enumeration of our main steps

We modify the first step in Ramadas' proof by working in the language of conformal blocks [25–27] and representation theory of Kac–Moody algebras. The differential form that we consider is a suitable correlation function. It plays a similar role as Ramadas' Harder–Narasimhan trace. (I would like to acknowledge this conceptual starting point given by Ramadas.) Roughly speaking, it corresponds to modifying the corresponding  $G$ -bundle (in the simplest Harder–Narasimhan stratum) around a finite set of additional points, and then taking a suitable mixed partial derivative

of the theta functions in the direction of these changes (the underlying principal bundle is actually trivial in this stratum). The modifications are controlled by the choice of simple positive roots of the Lie algebra for each of the additional points.

The correlation function is a differential form in the additional variables, with remarkable confluence properties as some of the additional points coalesce at the parabolic points, at a fixed point ( $\infty$  for us) and at a moving point. The works [25,26] (also see [8]) give local expressions for these correlation functions, allowing a bridge to the representation theory of Kac–Moody algebras (see Proposition 7.4).

The master function is the function discovered by Schechtman and Varchenko [23] (see Section 3.2). We prove the extension property of the differential form (the master function multiplied with the correlation function) to smooth compactifications by using the confluence properties alluded to in the previous paragraph. In particular, we prove more general forms of “vanishing along partial diagonals” by employing these techniques.

The crucial flatness assertion is proved by using results of Schechtman and Varchenko [23]. We use the interpretation [4] (also see [5, Section 1.8]) of the Schechtman–Varchenko formulas in terms of coinvariants of Kac–Moody algebras (also see [1] for a similar statement).

Exactly as in Ramadas’ paper, the last step (the fourth) involves applications of mixed Hodge theory.

## 2. Basic definitions in the theory of conformal blocks

We recall some definitions from Ueno’s book [27], which we will use as a reference for the theory of conformal blocks. Let  $\mathfrak{g}$  be a simple Lie algebra over  $\mathbb{C}$ . Fix the data of a Cartan decomposition of  $\mathfrak{g}$ :

$$\mathfrak{g} = \mathfrak{h} \oplus \sum_{\alpha \in \Delta} \mathfrak{g}_{\alpha}$$

where  $\Delta \subseteq \mathfrak{h}^*$  the set of roots is decomposed into a union  $\Delta_+ \cup \Delta_-$  of positive and negative roots. The set of simple (positive) roots is denoted by  $R$ .

A Killing form  $(\cdot, \cdot)$  on  $\mathfrak{g}$  induces one on  $\mathfrak{h}$  and  $\mathfrak{h}^*$ . Normalize the Killing form by requiring that  $(\theta, \theta) = 2$ , where  $\theta \in \Delta_+$  is the highest root.

### 2.1. Affine Lie algebras

The affine Lie algebra  $\hat{\mathfrak{g}}$  is defined to be

$$\hat{\mathfrak{g}} = \mathfrak{g} \otimes \mathbb{C}((\xi)) \oplus \mathbb{C}c,$$

where  $c$  is an element in the center of  $\hat{\mathfrak{g}}$  and the Lie algebra structure is defined by

$$[X \otimes f(\xi), Y \otimes g(\xi)] = [X, Y] \otimes f(\xi)g(\xi) + c(X, Y) \operatorname{Res}_{\xi=0}(g df),$$

where  $f, g \in \mathbb{C}((\xi))$  and  $X, Y \in \mathfrak{g}$ .

Introduce notation for the Lie subalgebras

$$\hat{\mathfrak{g}}_+ = \mathfrak{g} \otimes \mathbb{C}[[\xi]]\xi, \quad \hat{\mathfrak{g}}_- = \mathfrak{g} \otimes \mathbb{C}[[\xi^{-1}]]\xi^{-1}$$

so that we have a vector space decomposition

$$\hat{\mathfrak{g}} = \hat{\mathfrak{g}}_+ \oplus \mathfrak{g} \otimes \mathbb{C}c \oplus \hat{\mathfrak{g}}_-.$$

Let

$$X(n) = X \otimes \xi^n, \quad X = X(0) = X \otimes 1, \quad X \in \mathfrak{g}, \quad n \in \mathbb{Z}.$$

### 2.2. Representation theory of affine Lie algebras

Recall that finite dimensional irreducible representations of  $\mathfrak{g}$  are parameterized by the set of dominant integral weights  $P_+$  considered a subset of  $\mathfrak{h}^*$ . To  $\lambda \in P_+$ , the corresponding irreducible representation  $V_{\lambda}$  contains a non-zero vector  $v \in V_{\lambda}$  (the highest weight vector) such that

$$Hv = \lambda(H)v, \quad H \in \mathfrak{h},$$

$$X_\alpha v = 0, \quad X_\alpha \in \mathfrak{g}_\alpha, \quad \forall \alpha \in \Delta_+.$$

We will fix a level  $k$  in the sequel. Let  $P_k$  denote the set of dominant integral weights of level  $k$ . More precisely

$$P_k = \{\lambda \in P_+ \mid (\lambda, \theta) \leq k\},$$

where  $\theta$  is the highest (longest positive) root.

For each  $\lambda \in P_k$  there is a unique irreducible representation  $\mathcal{H}_\lambda$  of  $\hat{\mathfrak{g}}$  which satisfies the following properties:

- (1)  $V_\lambda = \{|v\rangle \mid \hat{\mathfrak{g}}_+ |v\rangle = 0\}$ .
- (2) The central element  $c$  of  $\hat{\mathfrak{g}}$  acts on  $\mathcal{H}_\lambda$  by multiplication by  $k$ .
- (3) Let  $|\lambda\rangle$  denote a highest weight vector in  $V_\lambda$ , then

$$X_\theta (-1)^{k-(\lambda, \theta)+1} |\lambda\rangle = 0,$$

where  $X_\theta$  is a non-zero element in the root space  $\mathfrak{g}_\theta$ . In fact  $\mathcal{H}_\lambda$  is generated by  $V_\lambda$  over  $\hat{\mathfrak{g}}$  with the above fundamental relation.

The representation  $\mathcal{H}_\lambda$  when  $\lambda = 0$  (still at level  $k$ ) is called the vacuum representation at level  $k$ .

### 2.3. Conformal blocks

We will work with conformal blocks on marked curves in  $\mathfrak{M}_{0,N}$  (that is, smooth curves of genus 0 with  $N$  marked points). But we will state the definitions in greater generality.

To define conformal blocks we will fix a stable  $N$ -pointed curve of genus  $g$  with formal neighborhoods  $\mathfrak{X} = (C; P_1, \dots, P_N, \eta_1, \dots, \eta_N)$ . Here we require:

- (1)  $C$  is smooth at the distinct points  $P_1, \dots, P_N$ .
- (2)  $C - \{P_1, \dots, P_N\}$  is affine.
- (3) A stability condition (finiteness of automorphisms of the pointed curve).
- (4) Isomorphisms  $\eta_i : \hat{\mathcal{O}}_{C, P_i} \xrightarrow{\sim} \mathbb{C}[[\xi_i]]$ ,  $i = 1, \dots, N$ .

Let  $\mathfrak{X}$  be as above, and choose  $\vec{\lambda} = (\lambda_1, \dots, \lambda_N) \in P_k^N$ . There are a number of definitions relevant to the situation: Let

$$\hat{\mathfrak{g}}_N = \bigoplus_{i=1}^N \mathfrak{g} \otimes_{\mathbb{C}} \mathbb{C}((\xi_i)) \oplus \mathbb{C}c$$

be the Lie algebra with  $c$  a central element and the Lie bracket given by

$$\left[ \sum_{i=1}^N X_i \otimes f_i, \sum_{i=1}^N Y_i \otimes g_i \right] = \sum_{i=1}^N [X_i, Y_i] \otimes f_i g_i + c \sum_{i=1}^N (X_i, Y_i) \text{Res}_{P_i}(g_i df_i).$$

Let

$$\mathfrak{g}(\mathfrak{X}) = \mathfrak{g} \otimes_{\mathbb{C}} \Gamma(C - \{P_1, \dots, P_N\}, \mathcal{O}) \hookrightarrow \hat{\mathfrak{g}}_N. \quad (2.1)$$

Let  $\vec{\lambda}$  be as above. Set

$$\mathcal{H}_{\vec{\lambda}} = \mathcal{H}_{\lambda_1} \otimes \dots \otimes \mathcal{H}_{\lambda_N}.$$

For a given  $X \in \mathfrak{g}$  and  $f \in \mathbb{C}((\xi_i))$ , define  $\rho_i(X \otimes f)$  an endomorphism of  $\mathcal{H}_{\vec{\lambda}}$  by

$$\rho_i(X \otimes f)|v_1\rangle \otimes \dots \otimes |v_N\rangle = |v_1\rangle \otimes \dots \otimes (X \otimes f|v_i\rangle) \otimes \dots \otimes |v_N\rangle$$

where  $|v_i\rangle \in \mathcal{H}_{\lambda_i}$  for each  $i$ .

We can now define the action of  $\hat{\mathfrak{g}}_N$  on  $\mathcal{H}_{\vec{\lambda}}$  by

$$(X_1 \otimes f_1, \dots, X_N \otimes f_N) |v_1\rangle \otimes \cdots \otimes |v_N\rangle = \sum_{i=1}^N \rho_i (X_i \otimes f_i) |v_1\rangle \otimes \cdots \otimes |v_N\rangle.$$

**Definition 2.1.** Define the space of conformal blocks

$$V_{\vec{\lambda}}^{\dagger}(\mathfrak{X}) = \text{Hom}_{\mathbb{C}}(\mathcal{H}_{\vec{\lambda}}/\mathfrak{g}(\mathfrak{X})\mathcal{H}_{\vec{\lambda}}, \mathbb{C}).$$

Define the set of dual conformal blocks,  $V_{\vec{\lambda}}(\mathfrak{X}) = \mathcal{H}_{\vec{\lambda}}/\mathfrak{g}(\mathfrak{X})\mathcal{H}_{\vec{\lambda}}$ . These are both finite dimensional  $\mathbb{C}$ -vector spaces which can be defined on families (and commute with base change) [27].

Following Dirac's bracket conventions, elements of  $V_{\vec{\lambda}}^{\dagger}(\mathfrak{X})$  (or  $\mathcal{H}_{\vec{\lambda}}^*$ ) are frequently denoted by  $\langle \Psi |$  and those of  $V_{\vec{\lambda}}(\mathfrak{X})$  (or of  $\mathcal{H}_{\vec{\lambda}}$ ) by  $|\Phi\rangle$  and the pairing by  $\langle \Psi | \Phi \rangle$ .

**Remark 2.2.** Let  $\langle \Psi | \in V_{\vec{\lambda}}^{\dagger}$ ,  $|\Phi\rangle \in \mathcal{H}_{\vec{\lambda}}$ ,  $X \in \mathfrak{g}$ , and  $f \in \Gamma(C - \{P_1, \dots, P_N\}, \mathcal{O})$ , then the following “gauge condition” holds (using the embedding (2.1))

$$\langle \Psi | X \otimes f | \Phi \rangle = 0.$$

## 2.4. Propagation of vacua

Add a new point  $P_{N+1}$  together with the vacuum representation  $V_0$  of level  $k$ , at  $P_{N+1}$ . Also fix a formal neighborhood at  $P_{N+1}$ . We therefore have a new pointed curve  $\mathfrak{X}'$ , and an extended  $\vec{\lambda}' = (\lambda_1, \dots, \lambda_N, \lambda_{N+1} = 0)$ . The propagation of vacuum gives an isomorphism

$$V_{\vec{\lambda}'}^{\dagger}(\mathfrak{X}') \xrightarrow{\sim} V_{\vec{\lambda}}^{\dagger}(\mathfrak{X}), \quad \langle \widehat{\Psi} | \mapsto \langle \Psi |$$

with the key formula

$$\langle \widehat{\Psi} | (|\Phi\rangle \otimes |0\rangle) = \langle \Psi | \Phi \rangle.$$

## 2.5. Correlation functions

Suppose  $\mathfrak{X} \in \mathfrak{M}_{g,N}$ . Let

$$\begin{aligned} \langle \Psi | &\in V_{\vec{\lambda}}^{\dagger}(\mathfrak{X}), & Q_1, \dots, Q_M &\in C - \{P_1, \dots, P_N\}, \\ |\Phi\rangle &\in \mathcal{H}_{\vec{\lambda}}, & Q_1, \dots, Q_M &\in C - \{P_1, \dots, P_N\}, & Q_i \neq Q_j, & i < j, \end{aligned}$$

and corresponding elements  $X_1, \dots, X_M \in \mathfrak{g}$ . There is a very important differential called a correlation function

$$\Omega = \langle \Psi | X_1(Q_1) X_2(Q_2) \cdots X_M(Q_M) | \Phi \rangle \in \bigotimes_{i=1}^M \Omega_{C, Q_i}^1.$$

Here  $\Omega_C^1$  is the vector bundle of holomorphic one-forms on  $C$ . One way to define  $\Omega$  is via propagation by vacua: add points  $Q_1, \dots, Q_M$  with formal coordinates  $\psi_1, \dots, \psi_M$  and consider the elements  $X_a(-1)|0\rangle$  in the vacuum representation at those points. Then

$$\Omega = \langle \widehat{\Psi} | X_1(-1)|0\rangle \otimes X_2(-1)|0\rangle \cdots X_M(-1)|0\rangle \otimes |\Phi\rangle d\psi_1 \cdots d\psi_M.$$

The differential form  $\Omega$  is independent of the chosen coordinates (see [27, p. 70] for more details).

### 3. Formulation of the extension theorem

We will henceforth consider the case  $C = \mathbb{P}^1$ , with a chosen  $\infty$  and a coordinate  $z$  on  $\mathbb{A}^1 = \mathbb{P}^1 - \{\infty\}$ . Consider distinct points  $P_1, \dots, P_N \in \mathbb{A}^1 \subset \mathbb{P}^1$  with  $z$ -coordinates  $z_1, \dots, z_N$  respectively. The standard coordinate  $z$  endows each  $P_i$  with a formal coordinate. Let  $\mathfrak{X}$  be the resulting  $N$ -pointed curve with formal coordinates.

**Definition 3.1.** For every positive root  $\delta$ , make a choice of a non-zero element  $f_\delta$  in  $\mathfrak{g}_{-\delta}$ .

Assume that we are given  $\lambda_1, \dots, \lambda_N \in P_k$ , such that  $\mu = \sum_{i=1}^N \lambda_i$  is in the root lattice. Write  $\mu = \sum n_p \alpha_p$ , where  $\alpha_p$  are the simple positive roots. It is easy to see that each  $n_p$  is non-negative (for example, by evaluating both sides on  $H_{\alpha_p}$ ).

Let  $|\vec{\lambda}\rangle = |\lambda_1\rangle \otimes \dots \otimes |\lambda_N\rangle$  be the product of the corresponding highest weight vectors. Now consider and fix a map  $\beta: [M] = \{1, \dots, M\} \rightarrow \Delta_+$ , so that  $\mu = \sum_{a=1}^M \beta(a)$ . The maximum value of  $M$  is  $M = \sum n_p$  and this is the main case. In such cases each  $\beta(a)$  is necessarily a simple positive root. However, in various arguments, we will need to consider smaller values of  $M$ .

Introduce variables  $t_1, \dots, t_M$  considered points on  $\mathbb{P}^1 - \{\infty, P_1, \dots, P_N\}$ . Consider, for every  $\langle \Psi | \in V_{\vec{\lambda}}^\dagger(\mathfrak{X})$ , the correlation function

$$\Omega = \Omega_\beta(\langle \Psi |) = \langle \Psi | f_{\beta(1)}(t_1) f_{\beta(2)}(t_2) \cdots f_{\beta(M)}(t_M) | \vec{\lambda} \rangle.$$

We will use the simplified notation  $\langle \Psi | f_\beta(\vec{t}) | \vec{\lambda} \rangle$  for the right-hand side of the above expression. Note that  $\Omega$  has at most poles of the first order along the sum  $D$  of the divisors,

- (a) for  $1 \leq a < b \leq M$  the divisor given by the closure of  $t_a = t_b$ ;
- (b) for  $a \in [M]$  and  $i \in [N]$ , the divisor  $t_a = z_i$ .

**Remark 3.2.** The correlation function  $\Omega$  is regular at the generic point of the divisor  $t_a = \infty$  for any  $a$ .

It is known that in the genus 0 situation, conformal blocks embed in the  $\mathfrak{g}$ -invariants in the dual of the tensor product  $(\otimes V_{\lambda_j})^*$  (see [27, Proposition 6.1]). The following is an amplification:

**Lemma 3.3.** The map  $V_{\vec{\lambda}}^\dagger(\mathfrak{X}) \rightarrow H^0((\mathbb{P}^1)^M, (\otimes_{a=1}^M p_a^* \Omega_{\mathbb{P}^1}^1)(D))$  given by  $\langle \Psi | \mapsto \Omega_\beta(\langle \Psi |)$ , is injective if  $M = \sum n_p$ .

**Proof.** Assume  $\Omega = 0$ . We can successively take the residues of the form  $\Omega$  in any of the variables. For example if we take the residue of  $\Omega$  about  $t_1 = z_1$ , we get a form

$$\langle \Psi | f_{\beta(2)}(t_2) \otimes \cdots \otimes f_{\beta(M)}(t_M) \mid (f_{\beta(1)} | \lambda_1) \rangle \otimes |\lambda_2\rangle \otimes \cdots \otimes |\lambda_N\rangle.$$

For every partition of  $[M]$  into  $N$  subsets  $I_1 \cup \dots \cup I_N$  and every enumeration of elements in  $I_j$

$$I_j = \{i_j(1), \dots, i_j(k_j)\},$$

we learn that

$$\langle \Psi | \prod_{j=1}^N f_{\beta(i_j(1))} f_{\beta(i_j(2))} \cdots f_{\beta(i_j(k_j))} | \lambda_j \rangle = 0.$$

Since the image of  $\langle \Psi |$  in the dual of the tensor product  $(\otimes V_{\lambda_j})^*$  is non-zero and invariant under the action of  $\mathfrak{g}$ , it follows that  $\langle \Psi | = 0$ : The weight spaces in the representations  $V_{\lambda_j}$  are obtained from the highest weight vectors by applying  $f_\alpha$ 's with  $\alpha$  simple, and we should only consider values of  $\langle \Psi |$  on vectors which are in the 0-weight space of  $\mathfrak{h}$ .  $\square$

### 3.1. Some properties of correlation functions

Using the gauge condition, it is possible to “simplify” correlation functions of the form  $\Omega = \langle \Psi | \prod_{a=1}^M f_{\beta(a)}(t_a) | \vec{\lambda} \rangle$ , so that a particular variable  $t_a$  no longer appears inside the correlation function:  $\Omega$  is  $dt_a \otimes$  the quantity

$$\sum_{j=1}^N \frac{1}{t_a - z_j} \langle \Psi | \prod_{b=1, b \neq a}^M f_{\beta(b)}(t_b) | \rho_j(f_{\beta(a)}) | \vec{\lambda} \rangle + \sum_{b=1, b \neq a}^M \frac{1}{t_a - t_b} \langle \Psi | [f_{\beta(a)}, f_{\beta(b)}](t_b) \prod_{\ell=1, \ell \neq a, \ell \neq b}^M f_{\beta(\ell)}(t_\ell) | \vec{\lambda} \rangle.$$

We have used the rule

$$dt_a \otimes dt_1 \cdots dt_{a-1} \cdots dt_{a+1} \cdots dt_M = dt_1 dt_2 \cdots dt_{a-1} dt_a \cdots dt_M. \quad (3.1)$$

Note that if  $\beta(a) + \beta(b)$  is not a root, then  $[f_{\beta(a)}, f_{\beta(b)}] = 0$ , and if it is a root, then  $[f_{\beta(a)}, f_{\beta(b)}]$  is equal to a non-zero multiple of  $f_{\beta(a)+\beta(b)}$  (note that we have picked these elements arbitrarily, see Definition 3.1).

**Remark 3.4.** We do not have summands with a polar term of order 2 in  $(t_a - t_b)$ . This is because  $(f_{\delta_1}, f_{\delta_2}) = 0$  for all positive roots  $\delta_1$  and  $\delta_2$ . Introduction of roots  $e_\delta$  will complicate the situation here, and introduce such polar terms (these will correspond to motions along the Harder–Narasimhan strata).

From the above considerations, it follows that  $\Omega$  as a form in  $t_a$  has at most simple pole at each of the  $z_i$ ’s and at the  $t_b$ ’s (if  $\beta(a) + \beta(b)$  is not a root then there is no pole at  $t_b$ ). We can iterate this procedure and obtain an expression for  $\Omega$  as a sum of terms, each of which has a simple denominator of total degree  $M$  which shows that  $\Omega$  is a log form in the sense of Hodge theory, see Section 8.4 (and Proposition 8.5) for a more refined statement. The refined statement will be used to compare  $\Omega$  with the Schechtman–Varchenko forms.

Some properties of  $\Omega$  are not easy to see from such an expression. The theory of [25–27] allows us to expand  $\Omega$  in a power series (on suitable angular sectors) as collections of the  $t$ -points come together (see Section 7.2 and Proposition 7.4).

### 3.2. The master function

Let  $\kappa = k + g^*$  where  $g^*$  is the dual Coxeter number of  $\mathfrak{g}$ . The following master function was discovered by Schechtman and Varchenko [23]:

$$\mathcal{R} = \prod_{1 \leq i < j \leq N} (z_i - z_j)^{-\frac{(\lambda_i, \lambda_j)}{\kappa}} \prod_{a=1}^M \prod_{j=1}^N (t_a - z_j)^{-\frac{(\lambda_j, \beta(a))}{\kappa}} \prod_{1 \leq a < b \leq M} (t_a - t_b)^{-\frac{(\beta(a), \beta(b))}{\kappa}}.$$

We will have occasion to use the master function even when  $M \neq \sum n_p$  (in such a case the  $\beta(a)$ ’s will be positive, but not necessarily simple roots). To analyze the situation when  $t_a$ ’s approach each other, it is convenient to have a definition of the master function even in the case  $\beta$  is defined only a subset  $A$  of  $[M]$  (with  $M = \sum n_p$ ). In this case the master function is a function of variables  $\{t_a, a \in A\}$  and

$$\mathcal{R} = \prod_{1 \leq i < j \leq N} (z_i - z_j)^{-\frac{(\lambda_i, \lambda_j)}{\kappa}} \prod_{a \in A} \prod_{j=1}^N (t_a - z_j)^{-\frac{(\lambda_j, \beta(a))}{\kappa}} \prod_{a, b \in A, a < b} (t_a - t_b)^{-\frac{(\beta(a), \beta(b))}{\kappa}}.$$

### 3.3. The extension theorem

Suppose  $M = \sum n_p$  (and hence  $\beta: [M] \rightarrow R \subseteq \Delta_+$ ). Let

$$X = \{(t_1, \dots, t_M) \in \mathbb{A}^M: t_a \neq t_b, 1 \leq a < b \leq M, t_a \neq z_i, i \in [N], a \in [M]\}.$$

Fix a sufficiently divisible positive integer  $C$  such that

$$C(\lambda_i, \lambda_j), C(\beta(a), \beta(b)), C(\beta(a), \lambda_i) \in \mathbb{Z}, \quad \forall a, b \in [M], i, j \in [N], a < b, i < j.$$

Consider an unramified (possibly disconnected) cover of  $X$  given by  $Y = \{(t_1, \dots, t_M, y) \mid y^{C_\kappa} = P\}$ , where

$$P = \prod_{1 \leq i < j \leq N} (z_i - z_j)^{-C(\lambda_i, \lambda_j)} \prod_{a=1}^M \prod_{j=1}^N (t_a - z_j)^{C(\lambda_j, \beta(a))} \prod_{1 \leq a < b \leq M} (t_a - t_b)^{-C(\beta(a), \beta(b))}. \quad (3.2)$$

Now fix  $\langle \Psi \mid \in V_{\lambda}^{\dagger}(\mathfrak{X})$  and set  $\Omega = \Omega_{\beta}(\langle \Psi \mid)$ . The following extension result holds:

**Theorem 3.5.**

- (1) *The multi-valued meromorphic form  $\mathcal{R}\Omega$  on  $X$  is square integrable.*
- (2) *The differential form  $p^*(\mathcal{R}\Omega)$  extends to an everywhere regular, single valued, differential form of the top order on any smooth and projective compactification  $\bar{Y} \supset Y$ .*

**Definition 3.6.** Let  $Z$  be an  $n$ -dimensional smooth algebraic variety, and  $\Gamma$  a possibly multi-valued  $n$ -form of the following form: For every  $p \in Z$ , there is an analytic open subset  $U$  of  $Z$  containing  $p$ , such that  $\Gamma$  can be expressed as  $\Gamma = f\omega$ , where

- (1)  $\omega$  is a (single valued) meromorphic form on  $U$ ;
- (2) some positive integer power of  $f$  is a (single valued) meromorphic function on  $U$ .

Let  $S \subset Z$  be an irreducible subvariety. We will denote the logarithmic degree of  $\Gamma$  along  $S$  by  $d^S(\Gamma)$ . Briefly: Blow up  $Z$  along  $S$ , and let  $E$  be the exceptional divisor. Then,  $d^S(\Gamma) - 1$  is the order of vanishing of (any branch of)  $\Gamma$  along  $E$ .

**Definition 3.7.** Let  $\tilde{D} \subset (\mathbb{P}^1)^M$  be the sum of the divisor  $D$  (defined before Lemma 3.3) and the divisors  $t_a = \infty$  (for  $a = 1, \dots, M$ ).

Ramadas' strategy [22] is to prove this kind of theorem by showing that the logarithmic degree of  $\mathcal{R}\Omega$  along any abnormal stratum of the divisor  $\tilde{D}$  is positive (see [19] for the notation that we use here). These abnormal strata are of three kinds:

- (S1) A certain subset of the  $t$ 's come together (to an arbitrary moving point). That is  $t_1 = t_2 = \dots = t_L$  after renumbering (possibly changing  $\beta$ ).
- (S2) A certain subset of the  $t$ 's come together to one of the  $z$ 's. That is  $t_1 = t_2 = \dots = t_L = z_1$  after renumbering (possibly changing  $\beta$ ).
- (S3) A certain subset of the  $t$ 's come together to  $\infty$ . That is  $t_1 = t_2 = \dots = t_L = \infty$  after renumbering (possibly changing  $\beta$ ).

More precisely we prove the following theorem which implies Theorem 3.5.

**Theorem 3.8.** Assume that  $M = \sum n_p$ . The logarithmic degree of  $\mathcal{R}\Omega$  along each of the strata (S1), (S2) and (S3) is positive.

The proof of Theorem 3.8 following the proof of a similar assertion in [22] will be broken up into three parts corresponding to the strata (S1), (S2) and (S3). The proofs of these three cases share some common features, chiefly that the degrees of the correlation functions  $\Omega$  are controlled by power series expansions (see Proposition 7.4), and also in that the poles along partial diagonals are seen as a reflection of the properties of the Lie algebra  $\mathfrak{g}$  (for example that the logarithmic degree along any strata of  $\Omega$  is  $\geq 0$ ). These three cases also differ in some important details. Therefore, we have given detailed proofs repeating some ideas and methods.



#### 4. Proof of Theorem 3.8 on the (S1) stratum: “When some of the points come together”

Let  $S$  be the set  $t_1 = \dots = t_L$ . Let  $B$  be the completion of the local ring of  $(\mathbb{P}^1)^M$  along  $S$ , at its generic point. Then  $B = K_S[[u_2, \dots, u_L]]$  where  $K_S$  is the function field of  $S$ , and  $u_i = t_i - t_1$ . Clear poles of  $\Omega$  at the generic point of  $S$  by multiplying by  $w = \prod_{1 \leq a < b \leq L} (t_a - t_b)$ :

$$w\Omega = g d\vec{t}, \quad g = \sum_{d \geq d_0} g_d(u_2, \dots, u_L).$$

Here  $g_d$  is a homogeneous polynomial in the  $u$ 's of (total) degree  $d$  with coefficients in  $K_S$ , and  $d_0$  is the smallest possible degree (so that  $g_{d_0} \neq 0$ ). Our task is to prove that the logarithmic degree of  $\mathcal{R}\Omega$  along  $S$ , written as  $d^S(\mathcal{R}\Omega)$ , satisfies the inequality

$$d^S(\mathcal{R}\Omega) = (d_0 - (L(L-1)/2) + (L-1)) - \sum_{1 \leq a < b \leq L} \frac{(\beta(a), \beta(b))}{\kappa} > 0 \quad (4.1)$$

(the term in the first bracket is the logarithmic degree of  $\Omega$ ).

##### 4.1. Reduction to a case in which “the lowest degree terms” have no poles as $t_a$ approaches $t_b$ for $1 \leq a < b \leq L$

We will try to reduce the number of variables. Let us suppose for example that

- $g_{d_0}$  is not divisible by  $u_2$ .

If  $\beta(1) + \beta(2)$  is not a positive root then  $\Omega$  does not have a pole along  $t_1 = t_2$ , so automatically  $g_{d_0}$  is divisible by  $u_2$ . So let us assume that  $\beta(1) + \beta(2)$  is a positive root  $\delta$ . Now divide  $g_d$  by  $u_2$  with remainder:

$$g_d = r_d(u_3, \dots, u_L) + u_2 q_d(u_2, \dots, u_L)$$

and  $r_{d_0} \neq 0$ .

Let us consider in parallel the new correlation function  $\Omega^*$  with variables  $t_1, t_3, \dots, t_M$  and  $\beta^*(1) = \delta$ , along the stratum  $S^*$  given by  $t_1 = t_3 = \dots = t_L$ . We multiply the corresponding correlation function by

$$w^* = \prod_{1 \leq a < b \leq L, a \neq 2, b \neq 2} (t_a - t_b)$$

and write an expansion

$$w^* \Omega^* = \sum_{d \geq \tilde{d}_0} \tilde{g}_d(u_3, \dots, u_L) dt_1 dt_3 \cdots dt_N.$$

We know that (see (3.1) and Section 3.1)

$$\Omega = dt_2 \otimes \frac{\Omega^*}{t_2 - t_1} + \widehat{\Omega}$$

where  $\widehat{\Omega}$  is regular at  $t_1 = t_2$  (and at most poles of the first order as  $t_1$  approaches the other variables). Multiply by  $w$ , and get

$$w\Omega = dt_2 \otimes w^* \prod_{a \geq 3} (t_2 - t_a) \Omega^* + w\widehat{\Omega}.$$

Set  $u_2 = 0$  (that is,  $t_2 = t_1$ ) now and get  $r_d(u_3, \dots, u_L) = -\tilde{g}_d(u_3, \dots, u_L) \prod_{a \geq 2} (t_1 - t_a)$ . We can therefore conclude that  $\tilde{g}_{d_0-(L-2)} \neq 0$ . Now if  $\tilde{g}_{d_0-(L-2)-s} \neq 0$ , then  $r_{d_0-s} \neq 0$  which would imply that  $g_{d_0-s} \neq 0$ . Therefore,

$$d_0 = \tilde{d}_0 + (L-2). \quad (4.2)$$

Therefore (where  $\mathcal{R}^*$  is the master function of  $\beta^*$  and the variables  $t_1, t_3, \dots, t_L$ )

$$\begin{aligned}
d^S(\beta, \mathcal{R}\Omega) - d^{S^*}(\beta^*, \mathcal{R}^*\Omega^*) \\
&= (L-2) + (L-1)(L-2)/2 - L(L-1)/2 + (L-1) - (L-2) - \frac{(\beta(1), \beta(2))}{\kappa} \\
&= -\frac{(\beta(1), \beta(2))}{\kappa}.
\end{aligned}$$

We obtain the equalities

$$d^S(\Omega) = d^{S^*}(\Omega^*), \quad (4.3)$$

$$d^S(\mathcal{R}\Omega) = d^{S^*}(\mathcal{R}^*\Omega^*) - \frac{(\beta(1), \beta(2))}{\kappa} \quad (4.4)$$

(think of this as “caused by” the loss of one variable  $t_2$ , a polar term  $(t_1 - t_2)^{-1}$ , and a fractional power  $(t_1 - t_2)^{-\frac{(\beta(1), \beta(2))}{\kappa}}$ ).

We continue this process as far as possible (we have not used the fact that  $\beta(a)$  are simple roots above, the same arguments apply even if  $\beta(a)$  are arbitrary positive roots). At every step “two variables come together”. We will remove the one with the larger subscript from our list of variables. So we will have a  $\tilde{\beta}$  defined on a subset  $A$  of  $[M]$ , a new correlation function  $\tilde{\Omega}$ , a new stratum  $\tilde{S}$  (of some points in  $\{t_a: a \in A\}$  equaling each other) and a new master function  $\tilde{\mathcal{R}}$  corresponding to  $\tilde{\beta}$  and the variables  $\{t_a\}$ ,  $a \in A$ . For every  $t \in A$ , let the set of  $a \in [M]$  that got together to give  $t$  be denoted by  $\text{Anc}(t)$  (“the set of maximal ancestors”). Let the set of points that descend from  $1, \dots, L$  be denoted by  $T$ . For example if the process ends at the first step as above, then  $A = \{1, 3, \dots, M\}$ ,  $T = \{1, 3, \dots, L\}$  and the ancestors of 1 are 1 and 2.

For convenience assume  $1 \in T$ . Note that  $A = T \cup \{L+1, \dots, M\}$ . The case  $|T| = 1$  is separately covered below.

**Remark 4.1.** At this point  $\tilde{\Omega}$  may still have poles along a partial diagonal  $t_a = t_b$ , where  $a, b \in T$ ,  $a \neq b$ . Our assumption is only that a suitable “lowest degree term” is pole-free.

The logarithmic degree of  $\tilde{\Omega}$  is at least  $|T| - 1$ , because the appropriate lowest degree term is divisible by all pairwise differences (the role of  $t_1$  in the above argument can be played by any of the variables  $t_1, \dots, t_L$ ):

$$d^{\tilde{S}}(\tilde{\Omega}) \geq |T| - 1. \quad (4.5)$$

Now, because of (4.4),

$$d^S(\mathcal{R}\Omega) = d^{\tilde{S}}(\tilde{\mathcal{R}}\tilde{\Omega}) - \sum_{t \in T} \sum_{a, b \in \text{Anc}(t), a < b} \frac{(\beta(a), \beta(b))}{\kappa} \quad (4.6)$$

we also have  $d^S(\Omega) = d^{\tilde{S}}(\tilde{\Omega})$  because of (4.3) and hence

$$d^S(\mathcal{R}\Omega) = d^{\tilde{S}}(\tilde{\Omega}) - \sum_{1 \leq a < b \leq L} \frac{(\beta(a), \beta(b))}{\kappa}. \quad (4.7)$$

Introduce  $\gamma = \sum_{a=1}^L \beta(a) = \sum_{t \in T} \tilde{\beta}(t)$ ,

$$(\gamma, \gamma) - \sum_{a=1}^L (\beta(a), \beta(a)) = 2 \sum_{1 \leq a < b \leq L} (\beta(a), \beta(b)) \quad (4.8)$$

so we find

$$d^S(\mathcal{R}\Omega) = d^{\tilde{S}}(\tilde{\Omega}) - \frac{(\gamma, \gamma)}{2\kappa} + \sum_{a=1}^L \frac{(\beta(a), \beta(a))}{2\kappa}. \quad (4.9)$$

We have two estimates for  $d^{\tilde{S}}(\tilde{\Omega})$ : it is at least  $|T| - 1$  and also at least  $\frac{(\gamma, \gamma)}{2\kappa} - 1$  (by Lemma 4.3). Assume that both estimates lead to lower bounds for  $d^S(\mathcal{R}\Omega)$ , which are  $\leq 0$ . The basic idea in the proof is that one gets both a lower bound and an upper bound for  $(\gamma, \gamma)$  which are in conflict.

We obtain,

$$|T| - 1 - \frac{(\gamma, \gamma)}{2\kappa} + \sum_{a=1}^L \frac{(\beta(a), \beta(a))}{2\kappa} \leq 0 \quad (4.10)$$

and

$$\frac{(\gamma, \gamma)}{2k} - 1 - \frac{(\gamma, \gamma)}{2\kappa} + \sum_{a=1}^L \frac{(\beta(a), \beta(a))}{2\kappa} \leq 0,$$

that is

$$\frac{g^*(\gamma, \gamma)}{2k\kappa} \leq 1 - \sum_{a=1}^L \frac{(\beta(a), \beta(a))}{2\kappa}$$

or that

$$\frac{(\gamma, \gamma)}{2\kappa} \leq \frac{k}{g^*} - \frac{k}{g^*} \sum_{a=1}^L \frac{(\beta(a), \beta(a))}{2\kappa}$$

which gives (using inequality (4.10)), the inequality

$$|T| \leq 1 + \frac{k}{g^*} - \left( \frac{k + g^*}{g^*} \right) \sum_{a=1}^L \frac{(\beta(a), \beta(a))}{2\kappa} = \frac{\kappa}{g^*} - \sum_{a=1}^L \frac{(\beta(a), \beta(a))}{2g^*}.$$

We will now cover the case  $|T| < \frac{\kappa}{g^*}$ . Use (4.6) and Lemma 4.2 to see that

$$d^S(\mathcal{R}\Omega) \geq d^{\tilde{S}}(\tilde{\mathcal{R}}\tilde{\Omega}) = |T| - 1 - \sum_{t, t' \in T, t < t'} \frac{(\tilde{\beta}(t), \tilde{\beta}(t'))}{\kappa}.$$

Using the numerical inequalities,  $(\tilde{\beta}(t), \tilde{\beta}(t')) \leq 2$ ,  $t, t' \in T$ , we find

$$d^S(\mathcal{R}\Omega) \geq |T| - 1 - \frac{|T|(|T| - 1)}{\kappa}.$$

If  $|T| < \kappa$  and  $|T| > 1$  then the above quantity is positive. If  $|T| = 1$ , then Lemma 4.2 below assures us that  $d^S(\mathcal{R}\Omega) > d^{\tilde{S}}(\tilde{\mathcal{R}}\tilde{\Omega}) = 0$ . We have thus covered all cases (we are assuming that  $L > 1$ ).

**Lemma 4.2.** For every positive root  $\delta = \sum_{i=1}^n \delta_i$  where  $\delta_i$  are positive simple roots (possibly repeated in the sum), and  $n > 1$ ,

$$I(\delta) = \sum_{1 \leq i < j \leq n} (\delta_i, \delta_j) < 0.$$

**Proof.** We divide the proof into two cases:

(1)  $\mathfrak{g} \neq G_2$ . If the theorem is true for  $\delta_1$  and  $\delta_2$  then the theorem is true for their sum  $\delta_1 + \delta_2$ , because

$$I(\delta_1 + \delta_2) = I(\delta_1) + I(\delta_2) + (\delta_1, \delta_2).$$

But since  $\delta_1, \delta_2$  are positive roots, whose sum is also a root, and  $\mathfrak{g} \neq G_2$ , we have the inequality  $(\delta_1, \delta_2) \leq 0$  (see [3, p. 278]). We can now induct and get the desired statement. At the first step,  $\delta_1$  and  $\delta_2$  are distinct simple roots whose sum is a root and  $I(\delta_1 + \delta_2) = (\delta_1, \delta_2) < 0$  (a strict inequality, see [9, Lecture 21]).

(2)  $\mathfrak{g} = G_2$ . This situation is easy enough for a direct verification. There are 6 positive roots:

$$\alpha_1, \quad \alpha_2, \quad \alpha_1 + \alpha_2, \quad 2\alpha_1 + \alpha_2, \quad 3\alpha_1 + \alpha_2, \quad 3\alpha_1 + 2\alpha_2,$$

also note (employing the normalization  $(\theta, \theta) = 2$ )

$$(\alpha_1, \alpha_1) = \frac{2}{3}, \quad (\alpha_2, \alpha_2) = 2, \quad (\alpha_1, \alpha_2) = -1. \quad \square$$

**Lemma 4.3.**  $d^{\tilde{S}}(\tilde{\Omega}) + 1 \geq \frac{1}{2k}(\gamma, \gamma)$ .

**Proof.** For ease in notation let us assume  $\beta = \tilde{\beta}$  and drop the assumption that  $\beta(a)$  are simple roots. We will now expand  $\Omega$  by a power series in  $u_2 = t_2 - t_1, \dots, u_L = t_L - t_1$ . To apply the considerations of Section 7.2 below, write (by propagation by vacuum introduce the vacuum representation at  $z_0 = t_1$  and consider the vector (where  $\xi_0 = z - t_1$ , so that  $u_i = \xi_0(t_i)$ )  $f_{\beta(1)}(-1)|0\rangle$  at that point):

$$\Omega = \sum_{b_2, \dots, b_L} u_2^{-b_2-1} \cdots u_L^{-b_L-1} \Omega_{\vec{b}} \quad (4.11)$$

in the angular sector  $0 < |u_L| < \cdots < |u_2| < \epsilon$  (with  $\epsilon$  depending upon  $t_1$ ) where  $\Omega_{\vec{b}}$  equals

$$\langle \Psi | f_{\beta(L+1)}(t_{L+1}) \cdots f_{\beta(M)}(t_M) \rho_0(f_{\beta(2)}(\xi_0^{b_2})) \cdots \rho_0(f_{\beta(L)}(\xi_0^{b_L})) | f_{\beta(1)}(-1)|0\rangle \otimes |\vec{\lambda}\rangle dt_1 du_2 \cdots du_L.$$

The logarithmic degree of the summand in (4.11) is  $-\sum_{a=2}^L b_a$ . Suppose that this summand is non-zero. Let  $\gamma = \sum_{a=1}^L \beta(a)$ , applying Proposition 7.4, we find,

$$-1 + \sum_{a=2}^L b_a \leq -\frac{(\gamma, \gamma)}{2k},$$

hence

$$-\sum_{a=2}^L b_a \geq \frac{(\gamma, \gamma)}{2k} - 1.$$

Note that  $d^S(\Omega)$ , the logarithmic degree of  $\Omega$ , is (at least) the minimum of  $-\sum_{a=2}^L b_a$  (such that  $\rho_0(f_{\beta(2)}(\xi_0^{b_2})) \cdots \rho_0(f_{\beta(L)}(\xi_0^{b_L})) | f_{\beta(1)}(-1)|0\rangle \neq 0$ ). The desired inequality follows (see Section 7.1).  $\square$

**Remark 4.4.** For  $\mathfrak{g} = \mathfrak{sl}_2$ , Lemma 4.3 implies Ramadas' vanishing theorem in the conformal block situation.

## 5. Proof of Theorem 3.8 for the (S2) stratum: “When some points come together at a finite parabolic point”

Let  $S$  be the set  $t_1 = \cdots = t_L = z_1$ . Let  $B$  be the completion of the local ring of  $(\mathbb{P}^1)^M$  along  $S$ , at its generic point. Then  $B = K_S[[u_1, u_2, \dots, u_L]]$  where  $K_S$  is the function field of  $S$ , and  $u_i = t_i - z_1$ . Clear poles of  $\Omega$  at the generic point of  $S$  by multiplying:

$$w = \prod_{1 \leq a < b \leq L} (t_a - t_b) \prod_{a=1}^L (t_a - z_1),$$

and then

$$w\Omega = g d\vec{t}, \quad g = \sum_{d \geq d_0} g_d(u_1, \dots, u_L),$$

where  $g_d$  is a homogeneous polynomial in the  $u$ 's of (total) degree  $d$  with coefficients in  $K_S$ , and  $d_0$  is the smallest possible degree (so that  $g_{d_0} \neq 0$ ). Our task is to prove that

$$d^S(\mathcal{R}\Omega) = (d_0 - (L(L-1)/2 + L) + L) - \sum_{1 \leq a < b \leq L} \frac{(\beta(a), \beta(b))}{\kappa} + \sum_{a=1}^L \frac{(\lambda_1, \beta(a))}{\kappa} > 0$$

which simplifies to

$$d^S(\mathcal{R}\Omega) = (d_0 - L(L-1)/2) - \sum_{1 \leq a < b \leq L} \frac{(\beta(a), \beta(b))}{\kappa} + \sum_{a=1}^L \frac{(\lambda_1, \beta(a))}{\kappa} > 0. \quad (5.1)$$

### 5.1. Reduction to a case in which “the lowest degree terms” have no poles as $t_a$ approaches $t_b$ for $1 \leq a < b \leq L$

We will reduce (by induction) to the case (but no longer requiring that  $\beta(a)$  are simple roots):  $g_{d_0}$  is divisible by  $(t_1 - t_2)$ , so that  $d_0 \geq L(L-1)/2$ . Let us suppose for example that  $g_{d_0}$  is not divisible by  $(t_1 - t_2)$ . If  $\beta(1) + \beta(2)$  is not a positive root then  $\Omega$  does not have a pole along  $t_1 = t_2$ , so automatically  $g_{d_0}$  is divisible by  $(t_1 - t_2)$ .

So let us assume that  $\beta(1) + \beta(2)$  is a positive root  $\delta$ . Now divide  $g_d$  by  $t_1 - t_2 = u_1 - u_2$  with remainder:

$$g_d = r_d(u_1, u_3, \dots, u_L) + (u_1 - u_2)q_d(u_1, \dots, u_L),$$

and  $r_{d_0} \neq 0$ .

Let us consider in parallel the new correlation function  $\Omega^*$  with variables  $t_1, t_3, \dots, t_M$  and  $\beta^*(1) = \delta$ , along the stratum  $S^*$  given by  $t_1 = t_3 = \dots = t_L = z_1$ . We multiply the corresponding correlation function by

$$w^* = \prod_{1 \leq a < b \leq L, a \neq 2, b \neq 2} (t_a - t_b) \prod_{a=1, a \neq 2}^L (t_a - z_1)$$

and write an expansion

$$w^* \Omega^* = \sum_{d \geq d_0} \tilde{g}_d(u_1, u_3, \dots, u_L) dt_1 dt_3 \cdots dt_N.$$

We know that

$$\Omega = dt_2 \otimes \frac{\Omega^*}{t_2 - t_1} + \widehat{\Omega}$$

where  $\widehat{\Omega}$  is regular at  $t_1 = t_2$ . Multiply by  $w$ , and get

$$w\Omega = dt_2 \otimes w^*(t_2 - z_1) \prod_{a \geq 3} (t_2 - t_a) \Omega^* + w\widehat{\Omega}.$$

Note that  $w\widehat{\Omega}$  does not have poles along the partial diagonals  $t_a = t_b$  where  $1 \leq a < b \leq L$ , and equal to 0 when  $u_2 = u_1$ . Set  $u_2 = u_1$  (that is,  $t_2 = t_1$ ) now and get  $r_d(u_1, u_3, \dots, u_L) = (t_1 - z_1) \prod_{a \geq 2} (t_1 - t_a) \tilde{g}_d(u_1, u_3, \dots, u_L)$ . We can therefore conclude that  $\tilde{g}_{d_0-(L-1)} \neq 0$ . Now if  $\tilde{g}_{d_0-(L-1)-s} \neq 0$ , then  $r_{d_0-s} \neq 0$  which would imply that  $g_{d_0-s} \neq 0$ . Therefore,

$$d_0 = \tilde{d}_0 + (L-1). \quad (5.2)$$

By a calculation similar to that of (4.4), we obtain  $d^S(\Omega) = d^{S^*}(\Omega^*)$ .

We continue this process until we obtain a situation where we cannot reduce any more. So we will have  $\tilde{\beta}, T \subseteq A, \tilde{\Omega}, \tilde{S}$ , etc., as in Section 4 (where  $\tilde{S}$  corresponds to points  $\{t_a: a \in T\}$  equaling  $z_1$ ). Let  $\gamma = \sum_{a=1}^L \beta(a) = \sum_{t \in T} \tilde{\beta}(t)$  (as before) and obtain

$$d^S(\mathcal{R}\Omega) = d^{\tilde{S}}(\tilde{\Omega}) + \frac{2(\lambda_1, \gamma) - (\gamma, \gamma)}{2\kappa} + \sum_{a=1}^L \frac{(\beta(a), \beta(a))}{2\kappa}. \quad (5.3)$$

We have two estimates for  $d^{\tilde{S}}(\tilde{\Omega})$ : it is at least 0 (because there no poles in the “smallest degree term” as  $t$ ’s approach each other, but there may be a first order pole in the “lowest degree term” as the  $t$ ’s approach  $z_1$ ) and also at least by Lemma 5.1,  $\frac{(\gamma, \gamma) - 2(\lambda_1, \gamma)}{2\kappa}$ . Assuming that both estimates lead to lower bounds for  $d^S(\mathcal{R}\Omega)$ , which are  $\leq 0$ , we conclude that

$$-\frac{(\gamma, \gamma) - 2(\lambda_1, \gamma)}{2\kappa} + \sum_{a=1}^L \frac{(\beta(a), \beta(a))}{2\kappa} \leq 0, \quad (5.4)$$

and

$$\frac{(\gamma, \gamma) - 2(\lambda_1, \gamma)}{2\kappa} - \frac{(\gamma, \gamma) - 2(\lambda_1, \gamma)}{2\kappa} + \sum_{a=1}^L \frac{(\beta(a), \beta(a))}{2\kappa} \leq 0,$$

that is

$$\frac{g^*((\gamma, \gamma) - 2(\lambda_1, \gamma))}{2k\kappa} \leq - \sum_{a=1}^L \frac{(\beta(a), \beta(a))}{2\kappa}. \quad (5.5)$$

Inequality (5.4) implies that the quantity  $(\gamma, \gamma) - 2(\lambda_1, \gamma)$  is  $> 0$ , while inequality (5.5) implies that it is  $< 0$ , a contradiction.

**Lemma 5.1.**  $d^{\tilde{\Omega}}(\tilde{\Omega}) \geq \frac{1}{2k}((\gamma, \gamma) - 2(\lambda_1, \gamma)).$

**Proof.** For ease in notation let us assume that  $\beta = \tilde{\beta}$  and drop the assumption that  $\beta(a)$  are simple roots. The proof parallels that of Lemma 4.3. Expand  $\Omega = \tilde{\Omega}$  by a power series in  $u_1 = t_1 - z_1, \dots, u_L = t_L - z_1$  (on a suitable angular sector). To apply the considerations of Section 7.2 below, write

$$\Omega = \sum_{b_1, \dots, b_L} u_1^{-b_1-1} \cdots u_L^{-b_L-1} \Omega_{\vec{b}}, \quad (5.6)$$

where  $\Omega_{\vec{b}}$  equals

$$\langle \Psi | f_{\beta(L+1)}(t_{L+1}) \cdots f_{\beta(M)}(t_M) \rho_1(f_{\beta(1)}(b_1)) \cdots \rho_1(f_{\beta(L)}(b_L)) | \vec{\lambda}_1 \rangle dt_1 dt_2 \cdots dt_L.$$

The logarithmic degree of the summand (assumed to be non-vanishing) in (5.6) is  $-\sum_{a=1}^L b_a$ . Let  $\gamma = \sum_{a=1}^L \beta(a)$ , applying Proposition 7.4, we find (where as above,  $d^S(\Omega)$  is the logarithmic degree of  $\Omega$ ) then

$$\sum_{a=1}^L b_a \leq \frac{2(\lambda_1, \gamma) - (\gamma, \gamma)}{2k},$$

and hence

$$d^S(\Omega) \geq \frac{(\gamma, \gamma) - 2(\lambda_1, \gamma)}{2k}. \quad \square$$

## 6. Proof of Theorem 3.8 for the (S3) stratum: “When some points come together at infinity”

Let  $S$  be the set  $t_1 = \cdots = t_L = \infty$ . Let  $B$  be the completion of the local ring of  $(\mathbb{P}^1)^M$  along  $S$ , at its generic point. Then  $B = K_S[[u_1, u_2, \dots, u_L]]$  where  $K_S$  is the function field of  $S$ , and  $u_i = \frac{1}{t_i}$ . We will use the coordinate  $u = \frac{1}{z}$  at infinity. Clear poles of  $\Omega$  at the generic point of  $S$  by multiplying:

$$\Omega' = w\Omega, \quad w = \prod_{1 \leq a < b \leq L} (u_a - u_b),$$

and then

$$\Omega' = g d\vec{u}, \quad g = \sum_{d \geq d_0} g_d(u_1, \dots, u_L)$$

where  $g_d$  is a homogeneous polynomial in the  $u$ 's of (total) degree  $d$  with coefficients in  $K_S$ , and  $d_0$  is the smallest possible degree (so that  $g_{d_0} \neq 0$ ). Note that the form  $\Omega$  is holomorphic in each of its variables at infinity. Our task is to prove

$$d^S(\mathcal{R}\Omega) = (d_0 - L(L-1)/2 + L) - \sum_{1 \leq a < b \leq L} \frac{(\beta(a), \beta(b))}{\kappa} - \sum_{a=1}^L \frac{(\beta(a), \beta(a))}{\kappa} > 0. \quad (6.1)$$

In the above, we have used the fact that the functions  $(t_a - t_b)$  and  $(t_a - z_i)$  have poles of order 1 at infinity in  $t_a$ , so the master function acquires an additional pole at infinity in each variable  $t_a$  of order  $\frac{1}{\kappa}$  times the quantity

$$\sum_{i=1}^N (\lambda_i, \beta(a)) - \sum_{b \neq a} (\beta(a), \beta(b)) = \left( \sum_{i=1}^N \lambda_i - \sum_{b=1}^M \beta(b), \beta(a) \right) + (\beta(a), \beta(a)) = (\beta(a), \beta(a)).$$

### 6.1. Reduction to a case in which “the lowest degree terms” have no poles as $t_a$ approaches $t_b$ for $1 \leq a < b \leq L$

Let us suppose for example that  $g_{d_0}$  is not divisible by  $(u_1 - u_2)$ . If  $\beta(1) + \beta(2)$  is not a positive root then  $\Omega$  does not have a pole along  $t_1 = t_2$ , so automatically  $g_{d_0}$  is divisible by  $(u_1 - u_2)$ .

So let us assume that  $\beta(1) + \beta(2)$  is a positive root  $\delta$ . Now divide  $g_d$  by  $(u_1 - u_2)$  with remainder:

$$g_d = r_d(u_2, \dots, u_L) + (u_1 - u_2)q_d(u_1, \dots, u_L),$$

and  $r_{d_0} \neq 0$ .

Let us consider in parallel the new correlation function  $\Omega^*$  with variables  $t_1, t_3, \dots, t_M$  and  $\beta^*(1) = \delta$ ,  $\beta^*(a) = \beta(a)$ ,  $a > 2$ , along the stratum  $S^*$  given by  $t_1 = t_3 = \dots = t_L = \infty$ . We multiply the corresponding correlation function by

$$w^* = \prod_{1 \leq a < b \leq L, a \neq 2, b \neq 2} (u_a - u_b)$$

and write an expansion

$$w^* \Omega^* = \sum_{d \geq \tilde{d}_0} \tilde{g}_d(u_1, u_3, \dots, u_L) du_1 du_3 \cdots du_N.$$

We know that

$$\Omega = du_2 \otimes \frac{\Omega^*}{u_2 - u_1} + \hat{\Omega}$$

where  $\hat{\Omega}$  is regular at  $u_1 = u_2$ . Multiply by  $w$ , and get

$$w\Omega = du_2 \otimes w^* \prod_{a \geq 3} (u_1 - u_a) \Omega^* + w\hat{\Omega}.$$

Set  $u_1 - u_2 = 0$  now and get  $r_d(u_1, u_3, \dots, u_L) = \tilde{g}_d(u_1, u_3, \dots, u_L) \prod_{a \geq 2} (u_1 - u_a)$ . We can therefore conclude that  $\tilde{g}_{d_0 - (L-2)} \neq 0$ . Now if  $\tilde{g}_{d_0 - (L-2) - s} \neq 0$ , then  $r_{d_0 - s} \neq 0$  which would imply that  $g_{d_0 - s} \neq 0$ . Therefore,

$$d_0 = \tilde{d}_0 + (L - 2). \quad (6.2)$$

We calculate,

$$d^S(\Omega) - d^{S^*}(\Omega^*) = (L - 2) + (L - 1)(L - 2)/2 - L(L - 1)/2 + 1 = 0,$$

and so,  $d^S(\Omega) = d^{S^*}(\Omega^*)$ .

We continue this process until we obtain a situation where we cannot reduce any more. So we will have  $\tilde{\beta}$ ,  $T \subseteq A$ ,  $\tilde{\Omega}$ ,  $\tilde{S}$ , etc., as in Section 4 (where  $\tilde{S}$  corresponds to points  $\{t_a : a \in T\}$  equaling infinity),

$$d^S(\mathcal{R}\Omega) = d^{\tilde{S}}(\tilde{\Omega}) - \sum_{1 \leq a < b \leq L} \frac{(\beta(a), \beta(b))}{\kappa} - \sum_{a=1}^L \frac{(\beta(a), \beta(a))}{\kappa}. \quad (6.3)$$

Recall from Section 4.1, the notation  $\text{Anc}(t)$ , for the set of “maximal ancestors” of  $t \in T$ . Introduce, as before  $\gamma = \sum_{a=1}^L \beta(a) = \sum_{t \in T} \tilde{\beta}(t)$ . Use (4.8) and the equality

$$\sum_{a=1}^L \frac{(\beta(a), \beta(a))}{2\kappa} = \sum_{t \in T} \sum_{a \in \text{Anc}(t)} \frac{(\beta(a), \beta(a))}{2\kappa}$$

to obtain the following expression:

$$d^S(\mathcal{R}\Omega) = d^{\tilde{S}}(\tilde{\Omega}) - \frac{(\gamma, \gamma)}{2\kappa} - \sum_{t \in T} \sum_{a \in \text{Anc}(t)} \frac{(\beta(a), \beta(a))}{2\kappa}. \quad (6.4)$$

We have two estimates for  $d^{\tilde{S}}(\tilde{\Omega}) = d^S(\Omega)$ : it is at least  $|T|$  (because the “lowest” degree term is divisible by all pairwise differences, compare with (4.5)) and also at least  $\frac{(\gamma, \gamma)}{2k}$  (this corresponds to  $\lambda_1 = 0$  in Lemma 5.1). Assuming that both estimates lead to lower bounds for  $d^S(\mathcal{R}\Omega)$ , which are  $\leq 0$ , we conclude that

$$|T| - \frac{(\gamma, \gamma)}{2\kappa} - \sum_{t \in T} \sum_{a \in \text{Anc}(t)} \frac{(\beta(a), \beta(a))}{2\kappa} \leq 0 \quad (6.5)$$

and

$$\frac{(\gamma, \gamma)}{2k} - \frac{(\gamma, \gamma)}{2\kappa} - \sum_{t \in T} \sum_{a \in \text{Anc}(t)} \frac{(\beta(a), \beta(a))}{2\kappa} \leq 0,$$

that is

$$\frac{g^*(\gamma, \gamma)}{2k\kappa} \leq \sum_{t \in T} \sum_{a \in \text{Anc}(t)} \frac{(\beta(a), \beta(a))}{2\kappa}$$

or that

$$\frac{(\gamma, \gamma)}{2\kappa} \leq \frac{k}{g^*} \sum_{t \in T} \sum_{a \in \text{Anc}(t)} \frac{(\beta(a), \beta(a))}{2\kappa}$$

which gives (using inequality (6.5)) the inequality

$$|T| \leq \left( \frac{k + g^*}{g^*} \right) \sum_{t \in T} \sum_{a \in \text{Anc}(t)} \frac{(\beta(a), \beta(a))}{2\kappa} = \sum_{t \in T} \sum_{a \in \text{Anc}(t)} \frac{(\beta(a), \beta(a))}{2g^*}.$$

It turns out that for every  $t \in T$ , the summand  $\sum_{a \in \text{Anc}(t)} \frac{(\beta(a), \beta(a))}{2g^*}$  on the right-hand side is  $< 1$ , leading to a contradiction. More precisely,

**Lemma 6.1.** *Let  $\delta$  be a positive root and  $\delta = \sum_{i=1}^s \delta_i$  where  $\delta_i$ ,  $1 = 1, \dots, s$ , are simple roots (possibly repeated). Then  $\sum_{i=1}^s (\delta_i, \delta_i) < 2g^*$ .*

**Proof.** The statement reduces to  $\delta = \theta$ , the highest root. Looking at the tables in [3], we can verify that if  $\theta = \sum b_i \delta_i$  where  $\delta_i$  are simple roots (without repetitions), then  $\sum b_i (\delta_i, \delta_i) = (g^* - 1)(\theta, \theta) = 2(g^* - 1)$ . (In the ADE case, the Coxeter number, which is the same as the dual Coxeter number, equals  $\sum b_i + 1$ . Also, in this case  $(\alpha, \alpha) = 2$  for all roots, so one has a simpler proof.)  $\square$

## 7. Power series expansions

### 7.1. Some generalities

Let  $f(z_1, \dots, z_n)$  be a meromorphic function defined in a neighborhood of the origin 0 in  $\mathbb{C}^n$ . Assume that  $\prod_i z_i \prod_{i < j} (z_i - z_j) f$  is holomorphic at the origin.

The multiplicity of  $f$  along a partial diagonal  $z_1 = \dots = z_s = 0$  can be calculated as follows. We can develop  $f$  in a power series as follows:

$$f = \frac{\prod_i z_i \prod_{i < j} (z_i - z_j) f}{\prod_i z_i \prod_{i < j} (z_i - z_j)},$$

and then develop the terms ( $1 \leq i < j \leq s$  on  $|z_i| < |z_j|$ )

$$\frac{1}{z_i - z_j} = -\frac{z_j^{-1}}{1 - (z_i/z_j)}$$

in power series in  $z_i/z_j$ .



Write

$$f(z_1, \dots, z_N) = \sum_{\vec{b}} z^{\vec{b}} g_{\vec{b}}(z_{s+1}, \dots, z_N),$$

where

- (1) the summation runs through vectors  $\vec{b} = (b_1, \dots, b_s) \in \mathbb{Z}^s$ ;
- (2)  $z^{\vec{b}} = z_1^{b_1} z_2^{b_2} \dots z_s^{b_s}$ ;
- (3)  $g_{\vec{b}}(z_{s+1}, \dots, z_N)$  is meromorphic with poles only along the partial diagonals;
- (4) the expansion is valid on  $0 < |z_1| < |z_2| < \dots < |z_s| < \epsilon$ ,  $|z_j - z_{j,0}| < \epsilon'$  for  $j = s+1, \dots, N$ , and some initial point  $(z_{s+1,0}, \dots, z_{N,0})$  (not on any partial diagonals) and positive real numbers  $\epsilon$  and  $\epsilon'$ .

Write  $|\vec{b}| = b_1 + b_2 + \dots + b_s$ . The following proposition is immediate.

**Proposition 7.1.** *The multiplicity of  $f$  along  $z_1 = \dots = z_s = 0$  is the minimum value of the set  $\{|\vec{b}| : g_{\vec{b}} \neq 0\}$ .*

## 7.2. Recollections from vertex operator algebra theory and Kac–Moody algebras

Conformal field theory gives an expansion of the correlation function about any point. The expansion of the correlation function (with notation as in Section 2.5)

$$\omega = \langle \Psi | X_1(Q_1) \cdots X_M(Q_M) | \Phi \rangle$$

about  $Q_1 = \dots = Q_L = P_1$  is, using the coordinate  $\xi_1$  (the answer is an element in  $\mathbb{C}((u_1))((u_2)) \cdots ((u_L))$ , where  $u_i = \xi_1(Q_i)$ ) an expression, valid in the region  $0 < |u_L| < |u_{L-1}| < \dots < |u_1| < \epsilon$  for some  $\epsilon$  (and other variables staying close to initial points), of the form

**Proposition 7.2.**

$$\Omega = \sum_{b_1, \dots, b_L} \omega_{\vec{b}} u_1^{-b_1-1} \cdots u_L^{-b_L-1}$$

with

$$\omega_{\vec{b}} = \langle \Psi | X_{L+1}(Q_{L+1}) \cdots X_M(Q_M) | \rho_1(X_1(\xi_1^{b_1})) \cdots \rho_L(X_L(\xi_1^{b_L})) | \Phi \rangle d\vec{u},$$

where  $d\vec{u} = du_1 \cdots du_L$ .

**Proof.** We know that multiplying  $\Omega$  by a (suitable power of)  $\prod_{i=1}^L u_i \prod_{1 \leq i < j \leq L} (u_i - u_j)$ , one gets a form holomorphic in  $(u_1, \dots, u_L)$  (in a neighborhood of  $0 \in \mathbb{C}^L$ ). So such an expansion of  $\Omega$  exists. To find the values of  $\omega_{\vec{b}}$ , we take iterated residues. More precisely by Theorem 3.24(3) in [27], we have

$$\begin{aligned} \text{Res}_{u_L=0} u_L^s \langle \Psi | X_1(Q_1) \cdots X_L(Q_L) X_{L+1}(Q_{L+1}) \cdots X_M(Q_M) | \Phi \rangle \\ = \langle \Psi | X_1(Q_1) \cdots X_{L-1}(Q_{L-1}) X_{L+1}(Q_{L+1}) \cdots X_M(Q_M) | \rho_1(X_L(\xi_1^s)) | \Phi \rangle, \end{aligned}$$

and we may iterate this procedure.  $\square$

**Remark 7.3.**

- (1) Note that if  $|\Phi\rangle = |\Phi_1\rangle \otimes \cdots \otimes |\Phi_N\rangle$ , then

$$\rho_1(X_1(\xi_1^{b_1})) \cdots \rho_L(X_L(\xi_1^{b_L})) |\Phi\rangle = (X_1(b_1) \cdots X_L(b_L) |\Phi_1\rangle) \otimes |\Phi_2\rangle \otimes \cdots \otimes |\Phi_N\rangle.$$

We may hence rewrite the expression for  $\omega$  as

$$\sum_{b_1, \dots, b_L} \langle \Psi | X_{L+1}(Q_{L+1}) \cdots X_M(Q_M) | (X_1(b_1) \cdots X_L(b_L) |\Phi_1\rangle) \otimes |\Phi_2\rangle \otimes \cdots \otimes |\Phi_N\rangle u_1^{-b_1-1} \cdots u_L^{-b_L-1} d\vec{u}.$$

- (2) The above expansion is valid for any choice of  $|\Phi_i\rangle \in \mathcal{H}_{\lambda_i}$  (not necessarily in  $V_{\lambda_i}$ ). In one of its applications (Lemma 4.3) we will propagate vacua and let  $P_1$  be the new point (after renaming the parabolic points), and consider expansions with  $|\Phi_1\rangle = f_\delta(-1)|0\rangle$  for a suitable positive root  $\delta$ .

### 7.3. Vanishing criteria

In a finite dimensional irreducible representation  $V_\lambda$  of a simple Lie algebra  $\mathfrak{g}$  with highest weight  $\lambda$ , one knows that for any weight  $\lambda'$  occurring in  $V_\lambda$ ,  $(\lambda', \lambda') \leq (\lambda, \lambda)$ . There is an analogous fact for integrable highest weight representations of Kac–Moody algebras (see [13, Chapters 7 and 12]). Using these results, we now formulate conditions under which an expression of the form  $X_1(b_1) \cdots X_L(b_L)|\lambda\rangle$ , where  $|\lambda\rangle$  is a highest weight vector in an integrable highest weight representation  $\mathcal{H}_\lambda$  of  $\hat{\mathfrak{g}}$  of level  $k$ , is necessarily zero.

**Proposition 7.4.** *Suppose that  $X_a = f_{\beta(a)} \in \mathfrak{g}_{-\beta(a)}$ , where  $\beta(a)$  are positive roots (not necessarily simple) and set  $\gamma = \sum_{a=1}^L \beta(a)$ . If  $X_1(b_1) \cdots X_L(b_L)|\lambda\rangle \neq 0$ , then*

$$\sum_{a=1}^L b_a \leq \frac{2(\lambda, \gamma) - (\gamma, \gamma)}{2k}. \quad (7.1)$$

### Remark 7.5.

- (1) In (7.1), the level  $k$  appears in the denominator (and not  $\kappa = k + g^*$ ). In the master functions that we consider, we have  $\kappa$  as a fractional exponent. The “small” difference between  $\kappa$  and  $k$  can be deemed responsible in part for the extension theorem (Theorem 3.8). In fact, the difference  $g^*$  is in a sense optimal (Lemma 6.1).  
 (2) In (7.2) and (7.1), inner products are computed in  $\mathfrak{h}^*$ .

### 7.4. Proof of Proposition 7.4

Proposition 7.4 is a direct corollary of Theorem 12.5, part (d) of [13]. To adjust to the notation there, we will include a few brief remarks. Extend the Lie algebra  $\hat{\mathfrak{g}}$  (as in [13]) by

$$\hat{\mathfrak{g}}' = \hat{\mathfrak{g}} + \mathbb{C}d, \quad \hat{\mathfrak{h}}' = \mathfrak{h} + \mathbb{C}c + \mathbb{C}d$$

with the commutations

$$[d, c] = 0, \quad [d, X(n)] = nX(n).$$

Extend a form  $\lambda \in \mathfrak{h}^*$  to  $\lambda \in \hat{\mathfrak{h}}'^*$  by setting  $\langle \lambda, \mathbb{C}c + \mathbb{C}d \rangle = 0$  where  $\langle \cdot, \cdot \rangle$  is the standard pairing of a vector space and its dual.

We define the elements  $\Lambda_0$  and  $\delta$  in the dual  $\hat{\mathfrak{h}}'^* = \mathfrak{h}^* + \mathbb{C}\Lambda_0 + \mathbb{C}\delta$  by  $\langle \delta, d \rangle = \langle \Lambda_0, c \rangle = 1$ ,  $\langle \delta, \mathfrak{h} + \mathbb{C}c \rangle = \langle \Lambda_0, \mathfrak{h} + \mathbb{C}d \rangle = 0$ . We extend the form  $(\cdot, \cdot)$  to  $\hat{\mathfrak{h}}'^*$  by putting

$$(\mathfrak{h}^*, \mathbb{C}\Lambda_0 + \mathbb{C}\delta) = (\delta, \delta) = (\Lambda_0, \Lambda_0) = 0, \quad (\delta, \Lambda_0) = 1.$$

Given a highest weight representation of  $\hat{\mathfrak{g}}$  of level  $k$ , we extend it to a representation of  $\hat{\mathfrak{g}}'$  by having  $d$  act on the highest weight vector by 0 (actually define  $d$  as  $-L_0 + \alpha \text{Id}$  for a suitable constant  $\alpha$  where  $L_0$  is obtained from the Sugawara tensor).

The highest weight of  $V_\lambda$  corresponds to the weight  $\lambda + k\Lambda_0 \in \hat{\mathfrak{h}}'^*$ , and  $X_1(b_1) \cdots X_L(b_L)|\lambda\rangle$  corresponds to the weight  $\lambda - \gamma + (\sum_{a=1}^L b_a)\delta + k\Lambda_0$ . Apply [13, Theorem 12.5, part (d)], to get the following inequality which implies inequality (7.1):

$$(\lambda - \gamma, \lambda - \gamma) + 2k \sum_{a=1}^L b_a \leq (\lambda, \lambda). \quad (7.2)$$

## 8. The map to a Gauss–Manin system

### 8.1. The KZ/WZW connection

Let us first recall the KZ connection on conformal blocks as expressed in [26] (see [27, Section 6.2]). Consider the tautological family of genus 0 pointed curves over the configuration space of  $N$  distinct points on  $\mathbb{A}^1$ ,  $\mathcal{C} = \mathbb{A}^N - \bigcup_{i \neq j} \{z = (z_1, \dots, z_N) \mid z_i = z_j\}$ . The pointed curve corresponding to  $z \in \mathcal{C}$  will be denoted by  $\mathfrak{X}(z)$ .

Consider a (local) family  $\langle \Psi \mid = \langle \Psi \mid (z) \in V_{\lambda}^{\dagger}(\mathfrak{X}(z))$  which we may view as a  $((V_{\lambda_1} \otimes \dots \otimes V_{\lambda_N})^*)^{\mathfrak{g}}$  valued function defined on a (small) analytic open subset  $U$  of  $\mathcal{C}$ . Such a family is flat if and only if, for any (constant, i.e. independent of  $z_1, \dots, z_N$ ) element  $|\vec{v}\rangle = |\nu_1\rangle \otimes \dots \otimes |\nu_N\rangle \in V_{\lambda_1} \otimes V_{\lambda_2} \otimes \dots \otimes V_{\lambda_N}$  and for each  $1 \leq i \leq n$ ,

$$\frac{d}{dz_i} \langle \Psi \mid \vec{v} \rangle = \frac{1}{\kappa} \sum_{j=1, j \neq i}^N \frac{\langle \Psi \mid \Omega_{ij} \mid \vec{v} \rangle}{z_i - z_j}, \quad (8.1)$$

where

$$\Omega_{ij} = \sum_{a=1}^{\dim \mathfrak{g}} \rho_i(J^a) \rho_j(J^a)$$

for any orthonormal basis  $\{J^a\}$  of  $\mathfrak{g}$ .

### 8.2. The map to cohomology

Now consider and fix a map  $\beta: [M] \rightarrow R$  (as before, where  $R$  is the set of simple roots), with  $\mu = \sum \lambda_i = \sum_{a=1}^M \beta(a)$  so that  $M = \sum n_p$ .

For  $z \in \mathcal{C}$ , let  $Y_z$  be the cover of

$$X_z = \{(t_1, \dots, t_M) \in \mathbb{A}^M: t_a \neq t_b, 1 \leq a < b \leq M, t_a \neq z_i, i = 1, \dots, N, a = 1, \dots, M\}$$

given by  $Y_z = \{(t_1, \dots, t_M, y) \mid y^{C_K} = P\}$ , where  $P$  and  $C$  were defined earlier (see (3.2)). The spaces  $Y_z$  organize into a smooth family of affine varieties  $\mathcal{Y} \rightarrow \mathcal{C}$ .

**Definition 8.1.** Define a map

$$S: V_{\lambda}^{\dagger}(\mathfrak{X}(z)) \rightarrow H^0(Y_z, \Omega_{Y_z}^M),$$

as follows: The map  $S$  takes  $\langle \Psi \mid (z) \in V_{\lambda}^{\dagger}(\mathfrak{X}(z))$  to the differential  $\mathcal{R}\Omega_{\beta}(\langle \Psi \mid)(z)$  (note that  $y = \mathcal{R}$  on  $Y_z$ ):

$$S(\langle \Psi \mid) = \mathcal{R}\Omega_{\beta}(\langle \Psi \mid (z)).$$

We recall that  $\Omega_{\beta}(\langle \Psi \mid (z))$  was constructed in Section 3, and the master function  $\mathcal{R}$  was defined in Section 3.2.

**Proposition 8.2.** Compose  $S$  with the evident morphism  $H^0(Y_z, \Omega_{Y_z}^M) \rightarrow H^M(Y_z, \mathbb{C})$ . The resulting map

$$T: V_{\lambda}^{\dagger}(\mathfrak{X}(z)) \rightarrow H^M(Y_z, \mathbb{C})$$

is a flat map (that is, preserves connections) as  $z$  varies (with the KZ/WZW connection on the left, and the Gauss–Manin connection on the right-hand side).

Proposition 8.2 follows from:

**Proposition 8.3.** For every (local) section

$$\langle \Psi \mid (z) \in V_{\lambda}^{\dagger}(\mathfrak{X}(z)),$$

there exists a relative algebraic  $M - 1$  form  $\omega$  on the fibers  $\mathcal{Y} \rightarrow \mathcal{C}$  (locally over  $\mathcal{C}$ ), such that for  $1 \leq i \leq N$ ,

$$\mathcal{R}\Omega_{\beta}(\nabla_{\frac{d}{dz_i}}(\langle \Psi |)) = \frac{d}{dz_i} \mathcal{R}\Omega_{\beta}(\langle \Psi |) + d(\omega),$$

where the differential operator  $d$  is the relative  $d$ -operator for the map  $\mathcal{Y} \rightarrow \mathcal{C}$ .

### 8.3. Proof of Proposition 8.3

The main point is that the form  $\mathcal{R}\Omega_{\beta}(\langle \Psi |)$  coincides in a suitable sense with a map in the work [23] for which a key flatness statement is proved there. Once this connection is made, the proposition follows immediately from results in [23].

Following [23] (where any complex value for  $\kappa$  is permitted), the element  $\langle \Psi | (z)$  produces an element in  $(\mathcal{M}^*)_0$  ( $\mathcal{M}$  is a tensor product of Verma modules, which surjects upon the tensor product  $\bigotimes_{i=1}^N V_{\lambda_i}$ , and  $(\mathcal{M}^*)_0$  is the zero-weight space of  $\mathcal{M}^*$ ).

A map is defined in [23]:

$$\eta_{\kappa} = \mathcal{M}_0^* \rightarrow \Omega^N(X_z, \mathcal{L}_{\lambda, \kappa}),$$

where  $\mathcal{L}_{\lambda, \kappa}$  is a rank one local system. We may take  $\mathcal{L}_{\lambda, \kappa}$  to be the trivial vector bundle with the connection  $d - d(\log \mathcal{R})$ . Locally, in the complex analytic topology we have a flat map  $\mathcal{L}_{\lambda, \kappa} \rightarrow (\mathcal{O}, \nabla = d)$  which takes 1 to  $\mathcal{R}^{-1}$ .

It is immediate from Proposition 8.5 and the definitions of the maps  $\eta_{\kappa}$  that (with our value for  $\kappa = k + g^*$ )

$$\eta_{-\kappa}(\langle \Psi |) = \mathcal{R}\Omega_{\beta}(\langle \Psi |). \quad (8.2)$$

**Remark 8.4.** Write  $\langle \Psi |$  in terms of the basis  $\{\delta(\gamma, \epsilon)\}$  in [23]:

$$\langle \Psi | = \sum_{\gamma, \epsilon} c_{\gamma, \epsilon} \delta(\gamma, \epsilon),$$

where  $c_{\gamma, \epsilon} = \langle \Psi | (f(\gamma, \epsilon))$ . The formula given for  $\eta(\delta(\gamma, \epsilon))$  should be compared with the coefficients in Proposition 8.5.

Proposition 8.3 now follows from [23, Theorems 7.2.5 and 7.2.5'] (compare with Eqs. (B.4) and (B.5) in [1]). Note that the KZ equation in [1] differs by a sign from ours. In [5] (see Theorems 1.8.5 and 1.8.6), a proof of the Schechtman–Varchenko formulas is given using coinvariants (dual to what we call correlation functions in this paper).

### 8.4. A formula for $\Omega$

Suppose  $M = \sum n_p$ ,  $\beta: [M] \rightarrow R$ ,  $\langle \Psi | \in V_{\lambda}^{\dagger}(\mathfrak{X})$  and  $\Omega = \Omega_{\beta}(\langle \Psi |)$ . The following proposition gives a formula for  $\Omega$  (see [5, Theorem 1.8.4], [1, Proposition 3.2] for similar statements, also see [1, (B.4) and (B.5)]).

**Proposition 8.5.**  $\Omega = \langle \Psi | w(t, z) \rangle dt_1 \cdots dt_M$ , where  $|w(t, z)\rangle \in V_{\lambda_1} \otimes \cdots \otimes V_{\lambda_N}$  is given by the formula

$$|w(t, z)\rangle = \sum_{\text{part}} \prod_{i=1}^N \left\langle \left( \prod_{a \in I_i} f_{\beta(a)}(t_a) \right) \middle| \lambda_i \right\rangle \quad (8.3)$$

with

$$\left\langle f_{\gamma_1}(u_1) f_{\gamma_2}(u_2) \cdots f_{\gamma_q}(u_q) \middle| \lambda_i \right\rangle = \sum_{\text{perm}} \frac{1}{(u_1 - u_2)(u_2 - u_3) \cdots (u_q - z_i)} (f_{\gamma_1} \cdots f_{\gamma_q} | v_i \rangle) \in V_{\lambda_i},$$

and where  $\sum_{\text{part}}$  stands for the summation over all partitions of  $I = \{1, \dots, M\}$  into  $N$  disjoint parts  $I = I_1 \cup I_2 \cup \cdots \cup I_N$  and  $\sum_{\text{perm}}$  the summation over all permutations of the elements of  $\{1, \dots, q\}$ .

**Proof.** We will not assume that  $M = \sum n_p$  (hence  $\beta(a)$  may be arbitrary positive roots), and proceed by induction. Fix vectors  $|v_i\rangle \in V_{\lambda_i}$ , not necessarily highest weight vectors, for  $i = 1, \dots, N$ . Let  $|\vec{v}\rangle = |v_1\rangle \otimes \dots \otimes |v_N\rangle$  and consider  $\Theta = \langle \Psi | f_{\beta(1)}(t_1) \cdots f_{\beta(M)}(t_M) | \vec{v} \rangle$ . We will prove by induction on  $M$  that

$$\Theta = \left( \sum_{part} \langle \Psi | \prod_{i=1}^N \left\langle \prod_{a \in I_i} f_{\beta(a)}(t_a) \right| v_i \right\rangle \right) d\vec{t} \quad (8.4)$$

where  $d\vec{t} = dt_1 \cdots dt_M$ .

The proof is by induction on  $M$ . If  $M = 1$ , then the result is clear: start with  $\Omega = \langle \Psi | f_{\beta}(t) | \vec{v} \rangle$ . Now use the function  $\frac{1}{z-t}$  and the gauge condition to write

$$\Theta = \sum_{i=1}^N \frac{1}{t_1 - z_i} \langle \Psi | \rho_i(f_{\beta}) | \vec{v} \rangle dt_1.$$

For  $M > 1$ , let us write  $\Theta = f_{\Theta}(t_1, \dots, t_M) d\vec{t}$ . We want to show that  $f_{\Theta}$  equals the right-hand side of (8.3) divided by  $d\vec{t}$  (we do this to get rid of the non-commuting  $dt_1, \dots, dt_M$ ). We will show that both sides of the desired equations are equal as functions of  $t_1$ . It is easy to see that both sides vanish at infinity. We need to show that they have equal polar parts at every finite point. Therefore, we need to analyze the behavior as

- (1)  $t_1$  approaches  $z_i$ . Let  $i = 1$  for simplicity. The polar part of  $\Theta$  is  $\frac{1}{t_1 - z_1} f_{\tilde{\Theta}}$  corresponding to a correlation function with variables  $t_2, \dots, t_M$  (same  $\beta$ 's) with  $|v_1\rangle$  changed to  $f_{\beta(1)}|v_1\rangle$ . On the right-hand side we need to consider only terms which have a fraction  $\frac{1}{t_1 - z_1}$ . A little thought convinces us that the equality of the polar parts at  $t_1 = z_1$  follows from induction.
- (2)  $t_1$  approaches  $t_a$ . In this case the polar part of  $f_{\Theta}$  is  $\frac{1}{t_1 - t_a} f_{\tilde{\Theta}}$  corresponding to a correlation function with points  $t_2, \dots, t_M$ , with  $f_{\beta(a)}$  replaced by  $[f_{\beta(1)}, f_{\beta(a)}]$  (a multiple of  $f_{\beta(a)+\beta(1)}$  if  $\beta(a) + \beta(1)$  is a root, zero otherwise). On the other side we should be looking at terms which have a  $t_1 - t_a$  or  $t_a - t_1$ . First all partitions considered should have  $t_1$  and  $t_a$  in the same part. So we are looking at words which have  $f_{\beta(1)}f_{\beta(a)}$  or  $f_{\beta(a)}f_{\beta(1)}$  as sub-words. We use the formula

$$a(u)b(v)f_u f_v - a(v)b(u)f_v f_u = a(v)b(v)[f_u, f_v] + O(u - v). \quad \square$$

**Remark 8.6.** The form (with notation as in Proposition 8.5)  $\mathcal{R}\langle \Psi | w(t, z) \rangle dt_1 \cdots dt_M (= \mathcal{R}\Omega_{\beta}(\langle \Psi |))$  is called a Schechtman–Varchenko form (these were introduced in [23]).

## 9. Unitarity

The remaining arguments for unitarity are exactly as in [22]. For completeness, we include these details. Let  $\overline{Y}_z$  be a smooth compactification of  $Y$ , which varies algebraically with  $z$ . That is, using resolution of singularities, find an embedding of  $\mathcal{C}$ -varieties  $\mathcal{Y} \subseteq \overline{\mathcal{Y}}$  so that  $\overline{\mathcal{Y}}$  is smooth and projective over  $\mathcal{C}$  and  $\overline{\mathcal{Y}} - \mathcal{Y}$  is a (relative) divisor with normal crossings.<sup>1</sup>

We therefore have a flat map

$$T: V_{\lambda}^{\dagger}(\mathfrak{X}(z)) \rightarrow H^M(Y_z, \mathbb{C})$$

which (because of Theorem 3.8) factors through an injective map (see Lemma 3.3)

$$\overline{T}: V_{\lambda}^{\dagger}(\mathfrak{X}(z)) \rightarrow H^0(\overline{Y}_z, \Omega^M) \subseteq H^M(\overline{Y}_z, \mathbb{C}).$$

(The second inclusion follows from Hodge theory.) It follows that

<sup>1</sup> We may have to shrink  $\mathcal{C}$ , but a flat unitary structure over a non-empty Zariski open subset of  $\mathcal{C}$  yields one over all of  $\mathcal{C}$ .

**Proposition 9.1.**

- (1)  $T : V_{\lambda}^{\dagger}(\mathfrak{X}(z)) \rightarrow H^M(Y_z, \mathbb{C})$  is injective (and flat for connections by [23], as stated in Proposition 8.2).
- (2)  $\bar{T} : V_{\lambda}^{\dagger}(\mathfrak{X}(z)) \rightarrow H^M(\bar{Y}_z, \mathbb{C})$  is flat for the connections (KZ on one side and Gauss–Manin on the other).

The first part uses results of Deligne [6]; and the second uses the first, and a semi-simplicity theorem in (pure) Hodge theory. The first part is a theorem of Varchenko for  $\mathfrak{g} = \mathfrak{sl}_2$  (see [28, Theorem 14.6.4]).

The Hodge form on  $H^M(\bar{Y}_z, \mathbb{C})$  is given on de Rham classes by  $|\omega|^2 = (\sqrt{-1})^M \int_{\bar{Y}_z} \omega \wedge \bar{\omega}$ . The restriction of Hodge form to  $H^0(\bar{Y}_z, \Omega^M) \subset H^M(\bar{Y}_z, \mathbb{C})$  gives a unitary metric.

It is known that the Hodge form is induced by the cup product on cohomology, and is therefore preserved by the Gauss–Manin connection (note that this connection, in general may not preserve  $H^0(\bar{Y}_z, \Omega^M)$ ). It then follows that we can induce the desired unitary metric on  $V_{\lambda}^{\dagger}(\mathfrak{X}(z))$  through the map  $\bar{T}$ . Theorem 1.1 therefore holds.

**Remark 9.2.** As in [22], the image of  $\bar{T}$ , lands inside a weight space (for a suitable character) of  $H^0(\bar{Y}_z, \Omega^M)$ . It would be interesting to obtain a characterization of the image of  $\bar{T}$  (compare with Theorem 7 in [20] for the larger KZ system).

### 9.1. An explicit form of the metric

Given a conformal block  $\langle \Psi |$ , let  $\pi(\langle \Psi |) = \mathcal{R}\Omega(\langle \Psi |)$  be the corresponding Schechtman–Varchenko form. Then, the (KZ-invariant) unitary metric (up to a constant  $C\kappa$ ) is given by the convergent integral (as conjectured in [7]):

$$|\langle \Psi ||^2 = (\sqrt{-1})^M \int_{(\mathbb{P}^1)^M} \pi(\langle \Psi |) \wedge \overline{\pi(\langle \Psi |)}.$$

Note that the integrand on  $(\mathbb{P}^1)^M$  is single valued (roots of unity have norm one).

## 10. Configuration spaces and moduli spaces

It is easy to see that  $\mathfrak{M} = \mathfrak{M}_{0,N}$  is a smooth affine variety by identifying  $\mathfrak{M}$  with the configuration space of  $N - 3$  points on  $\mathbb{A}^1$ . It is known that there is a bundle of conformal blocks  $\mathcal{V}$  on  $\mathfrak{M}$  which comes equipped with a flat connection  $\nabla_{\mathfrak{M}}$  induced from the connection on conformal block bundles on configuration spaces (from Section 8.1).

**Remark 10.1.**

- (1) Define two unitary metrics on a complex vector space  $V$  to be projectively equivalent if they are positive real multiples of each other. A projective metric on  $V$  is an equivalence class under this relation.
- (2) Fix an actual unitary metric on  $V$ . There is a bijection between elements of  $\mathrm{PGL}(V)$  that carry this metric on  $V$  to a metric projectively equivalent to it, and elements of  $\mathrm{PU}(V)$ . Here  $\mathrm{PGL}(V) = \mathrm{GL}(V)/\mathbb{C}^*$ ,  $\mathrm{PU}(V) = \mathrm{U}(V)/\mathrm{U}(1)$ , and  $\mathrm{U}(1) = \{z \in \mathbb{C} \mid z\bar{z} = 1\}$ .

Consider the configuration space  $\mathcal{C}$  as in Section 8.1. There is a natural map  $\pi : \mathcal{C} \rightarrow \mathfrak{M}$ . The bundle of conformal blocks over  $\mathcal{C}$  considered in this paper is canonically equal to  $\pi^*\mathcal{V}$ . We have constructed a unitary metric  $(\cdot, \cdot)$  on  $\pi^*\mathcal{V}$  which is preserved by the connection  $\nabla_{\mathcal{C}}$  (as in Section 8.1).

The connection  $\nabla_{\mathcal{C}}$  on  $\pi^*\mathcal{V}$  is equal to  $\pi^*\nabla_{\mathfrak{M}}$  only as a projective connection (recall that the theory in [26] produces a well defined flat projective connection). Therefore, locally on  $\mathcal{C}$ ,  $\nabla_{\mathcal{C}} - \pi^*\nabla_{\mathfrak{M}} = df \otimes 1_{\pi^*\mathcal{V}}$  for a local function  $f$  on  $\mathcal{C}$ . If  $v$  is a local section of  $\mathcal{V}$  on  $\mathcal{M}$  and  $X$  a tangent vector field on the fibers of  $\mathcal{C} \rightarrow \mathcal{M}$ , we find  $\nabla_{\mathcal{C},X}\pi^*v = (\pi^*\nabla_{\mathfrak{M}})_X v + X(f)v = X(f)v$  and so,

$$X(\pi^*v, \pi^*v) = (\nabla_{\mathcal{C},X}\pi^*v, \pi^*v) + (\pi^*v, \nabla_{\mathcal{C},X}\pi^*v) = (X(f) + \overline{X(f)})(\pi^*v, \pi^*v).$$

The above argument shows that for  $x, y \in \mathcal{C}$  with  $\pi(x) = \pi(y) = p$ , the two metrics on  $\pi^*\mathcal{V}_x = \pi^*\mathcal{V}_y = \mathcal{V}_p$  are projectively equivalent. Hence the projective monodromy group of  $(\mathcal{V}, \nabla)$  on  $\mathfrak{M}$  about a base point  $p \in \mathfrak{M}$  is contained in  $\mathrm{PU}(\mathcal{V}_p)$  where the metric on  $\mathcal{V}_p$  is any element in the projective equivalence class constructed above on  $\mathcal{V}_p$ .

## Acknowledgements

This paper builds upon the beautiful recent work of T.R. Ramadas [22]. In addition, the work of E. Looijenga [19] and A. Varchenko (unpublished) were helpful to us. Some ideas in the alternate proof of Ramadas' vanishing theorem in these works where variables are set equal to each other, and in the pole analysis at finite parabolic points in the extension theorem, were used in the proof of Theorem 3.8.

I am deeply indebted to Najmuddin Fakhruddin for an extensive communication on the subject, his comments, clarifications and corrections, constant encouragement, and for interest in this work. I thank Madhav Nori for bringing Ramadas' work to my attention (in 2004), and for many conversations over the years on related questions. I thank A. Varchenko for pointing out many references (in particular, the works of Kirillov and Wenzl), and for his comments on an earlier version. I thank I. Cherednik, A. Kirillov, S. Kumar, E. Looijenga and H. Wenzl for useful discussions and correspondence.

## References

- [1] H. Awata, A. Tsuchiya, Y. Yamada, Integral formulas for the WZNW correlation functions, *Nuclear Phys. B* 365 (3) (1991) 680–696.
- [2] H. Babujian, R. Flume, Off-shell Bethe ansatz equation for Gaudin magnets and solutions of Knizhnik–Zamolodchikov equations, *Modern Phys. Lett. A* 9 (22) (1994) 2029–2039.
- [3] N. Bourbaki, *Groupe et algèbres de Lie*, chapitres 4, 5 et 6, Masson, Paris, 1981.
- [4] I. Cherednik, Integral solutions of trigonometric Knizhnik–Zamolodchikov equations and Kac–Moody algebras, *Publ. RIMS* 27 (1991) 727–744.
- [5] I. Cherednik, *Double Affine Hecke Algebras*, London Mathematical Society Lecture Notes Series, vol. 319, 2005.
- [6] P. Deligne, *Théorie de Hodge II*, *Inst. Hautes Études Sci. Publ. Math.* 40 (1971) 5–57.
- [7] F. Falceto, K. Gawedzki, A. Kupiainen, Scalar product of current blocks in WZW theory, *Phys. Lett. B* 260 (1–2) (1991) 101–108.
- [8] E. Frenkel, D. Ben-Zvi, *Vertex Algebras and Algebraic Curves*, second ed., *Mathematical Surveys and Monographs*, vol. 88, Amer. Math. Soc., 2004.
- [9] W. Fulton, J. Harris, *Representation Theory: A First Course*, *Graduate Texts in Mathematics/Readings in Mathematics*, vol. 129, Springer-Verlag, New York, 1991.
- [10] K. Gawedzki, A. Kupiainen,  $SU(2)$  Chern–Simons theory at genus zero, *Comm. Math. Phys.* 135 (1991) 531–546.
- [11] K. Gawedzki, *Lectures on conformal field theory*, in: *Quantum Fields and Strings: A Course for Mathematicians*, vols. 1, 2 (Princeton, NJ, 1996–1997), Amer. Math. Soc., Providence, RI, 1999, pp. 727–805.
- [12] N. Hitchin, Flat connections and geometric quantization, *Comm. Math. Phys.* 131 (1994) 347–380.
- [13] V. Kac, *Infinite Dimensional Lie Algebras*, third ed., Cambridge University Press, Cambridge, UK, 1990.
- [14] D. Kazhdan, G. Lusztig, Tensor categories arising from affine Lie algebras. I–IV, *J. Amer. Math. Soc.* 6 (1993) 905–947, 949–1011, *J. Amer. Math. Soc.* 7 (1994) 335–381, 383–453.
- [15] A. Kirillov, On an inner product in modular tensor categories, *J. Amer. Math. Soc.* 9 (4) (1996) 1135–1169.
- [16] A. Kirillov, On inner product in modular tensor categories. II. Inner product on conformal blocks and affine inner product identities, *Adv. Theor. Math. Phys.* 2 (1) (1998) 155–180.
- [17] A. Kirillov, H. Wenzl, Electronic correspondence with the author, February 2011.
- [18] E. Looijenga, From WZW models to modular functors, *arXiv:1009.2245*.
- [19] E. Looijenga, Unitarity of  $SL(2)$ -conformal blocks in genus zero, *J. Geom. Phys.* 59 (2009) 654–662.
- [20] E. Looijenga, A topological interpretation of the KZ system, in: *Proceedings of the Second MSJ Seasonal Institute 2009 on Arrangements of Hyperplanes*, in press, *arXiv:1003.2033v6*.
- [21] E. Mukhin, V. Tarasov, A. Varchenko, Schubert calculus and representations of the general linear group, *J. Amer. Math. Soc.* 22 (4) (2009) 909–940.
- [22] T.R. Ramadas, The ‘Harder–Narasimhan trace’ and unitarity of the KZ/Hitchin connection: genus 0, *Ann. of Math.* 169 (1) (2009) 1–39.
- [23] V.V. Schechtman, A.N. Varchenko, Arrangements of hyperplanes and Lie algebra homology, *Invent. Math.* 106 (1991) 139–194.
- [24] C. Sorger, La formule de Verlinde, in: *Séminaire Bourbaki*, Exp. 794, 1994.
- [25] A. Tsuchiya, Y. Kanie, Vertex operators in conformal field theory on  $\mathbb{P}^1$  and monodromy representations of braid group, in: *Conformal Field Theory and Solvable Lattice Models*, Kyoto, 1986, in: *Adv. Stud. Pure Math.*, vol. 16, Academic Press, Boston, MA, 1988, pp. 297–372.
- [26] A. Tsuchiya, K. Ueno, Y. Yamada, Conformal field theory on universal family of stable curves with gauge symmetries, in: *Integrable Systems in Quantum Field Theory and Statistical Mechanics*, in: *Adv. Stud. Pure Math.*, vol. 19, Academic Press, Boston, MA, 1989, pp. 459–566.
- [27] K. Ueno, *Conformal Field Theory with Gauge Symmetry*, *Fields Institute Monographs*, vol. 24, Amer. Math. Soc., Providence, RI, 2008.
- [28] A. Varchenko, *Multidimensional Hypergeometric Functions and Representation Theory of Lie Algebras and Quantum Groups*, *Advanced Series in Mathematical Physics*, vol. 21, World Scientific Publishing Co., River Edge, NJ, 1995.
- [29] H. Wenzl,  $C^*$  tensor categories from quantum groups, *J. Amer. Math. Soc.* 11 (1998) 261–282.